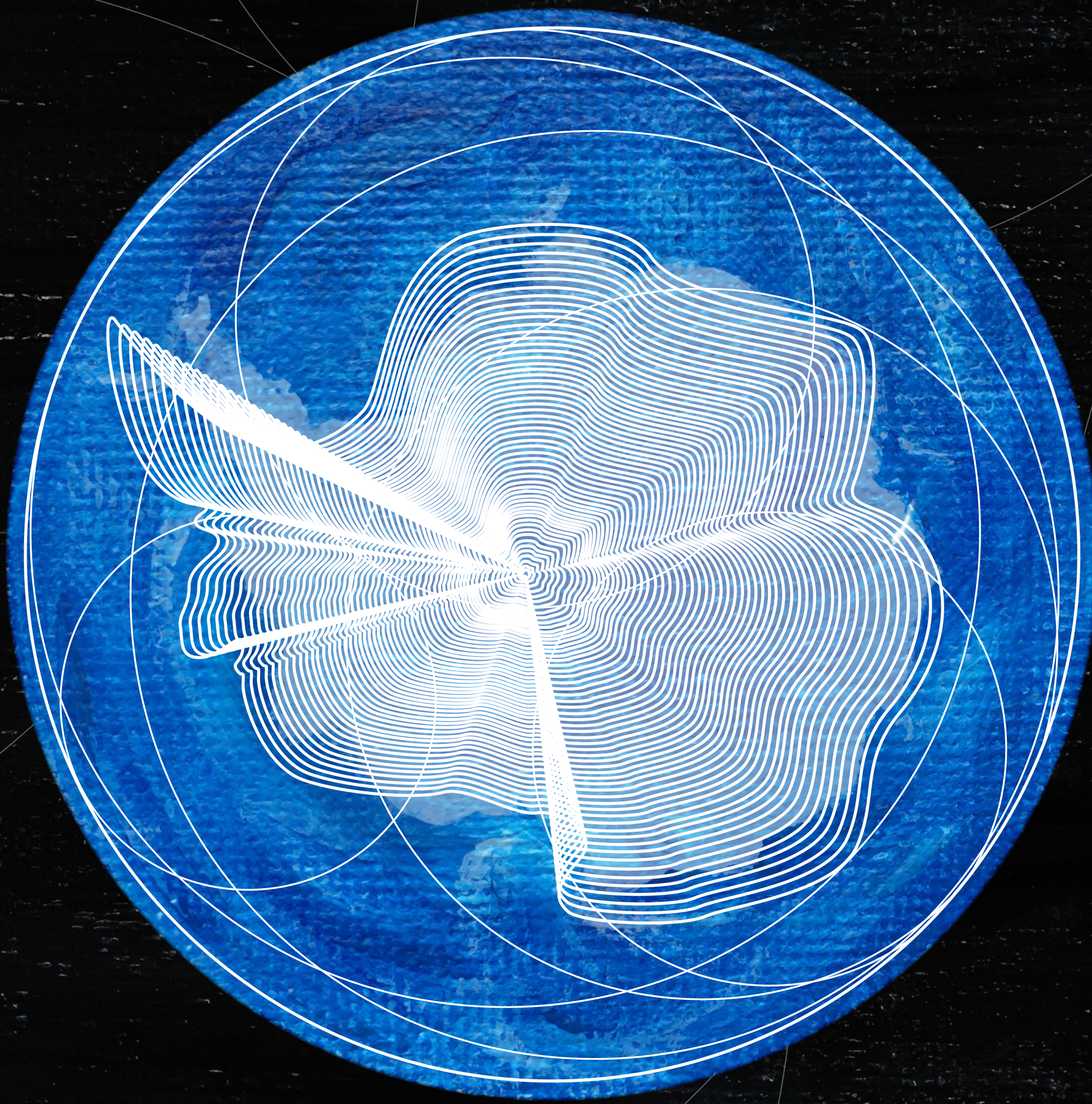
MSURJ

MCGILL SCIENCE UNDERGRADUATE RESEARCH JOURNAL
VOLUME 13 \ ISSUE 1 \ MARCH 2018



ON THE COVER

Climate modelling allows us to forecast changes in global climate patterns while allowing us to better understand past events. Current models predict that Antarctic ice sheets should have decreased in the time period between 1975 and 2017. But the opposite has happened.

On page 34, Sauvé et al. examine the role of natural variability and anthropogenic forcing on climate models that could explain Antarctic ice trends. Pictured on the cover is an artistic rendition of the Antarctic ice sheets.

MCGILL SCIENCE
UNDERGRADUATE
RESEARCH JOURNAL

BOARD 2017-2018

805 Sherbrooke St. West, Room 1B23
Montreal, Quebec, H3A 2K6 Canada
Phone: (514) 398-6979
Fax: (514) 398-6766
Email: msurj.sus@mail.mcgill.ca
Website: msurj.com

TABLE OF CONTENTS

FOREWORD

Dear Reader,

In 1983, a McGill Computer Science student, Alan Emtage, invented the search engine. This single invention revolutionised how we interact with information. Long gone are the days of painstakingly labouring through library directories. Ask a young scientist when they last held a physical research journal in their hands; the answer will likely be never.

We thus present to you the first completely paperless volume of the McGill Science Undergraduate Research Journal. With Volume 13, we remain devoted to upholding our mandate of promoting undergraduate scientific research regardless of academic affiliation. And with our blog, the Abstract, we continue to make strides in making scientific research more accessible to the general public.

Reflected on this journal is the same spirit of ingenuity and innovation that drove Alan Emtage thirty-five years ago. This year's selection spans the range of disciplines from earth science to biology to nanochemistry to astrophysics. Regardless of your background, we hope you find something of interest in the pages that follow.

Nick Pant & Rachel Greben
Editors-in-Chief

ACKNOWLEDGEMENTS

The McGill Science Undergraduate Research Journal would like to thank its generous contributors without whom this journal would not have been possible.

We thank Dean Bruce Lennox who played a pivotal role in helping MSURJ set new goals for the upcoming decade; without his advice and expertise, MSURJ would not be the success it is currently. We would also like to thank the McGill Librarians for their help in transitioning MSURJ to an online electronic system. Lastly, we would like to acknowledge Mr. Victor Chisholm for his continuous support and guidance.

We thank all the donors in the McGill community for their generous support:

Faculty of Science
Faculty of Medicine
Department of Biochemistry
Chemistry Undergraduate Student Society (CUSS)
Pharmacology Integrative League of Students (PILS)
Physiology Undergraduate League of Students (PULS)
Biochemistry Undergraduate Society (BUGS)
Microbiology and Immunology Students Association (MISA)
Science Undergraduate Society of McGIll (SUS)

We extend our gratitude to the numerous peer reviewers who took the time to review all of the submissions. Lastly, we wish to recognize the student authors whose efforts have made the journal possible.

Submitted: 01/14/18

Research Article

[1]Department of Physics, McGill University, Montreal, QC, Canada

Keywords

Astrophysics, High-Energy, Pulsar, Gamma-Ray, *Fermi*-LAT

Email Correspondence

amanda.cook@mail.mcgill.ca

Amanda Cook[1]

Exploration of *Fermi*-LAT Data: An Analysis of Pulsar J1930+1852

Abstract

Background: *Fermi*-LAT's 9-year data set of astrophysical gamma-rays (recently reprocessed) has revealed many new astrophysical sources. A closer analysis of one of these previously unseen sources, PSR J1930+1852 and associated pulsar wind nebula, G54.0+0.3, could help to confirm the gamma-ray emission mechanism of pulsars.

Methods: An investigation and analysis of PSR J1930+1852 and PWN G54.0+0.3 using *Fermi*-LAT data and science tools using maximum likelihood fitting is detailed.

Results: A 4.3 σ (p = 0.000017) excess above background was observed at the coordinates of the pulsar/pulsar wind nebula and the sources spectrum appears to be consistent with a single power law.

Limitations: The sources in the models are modelled as point sources. Further studies may want to consider the possibility of extended sources in the modelled region.

Conclusion: There is evidence for a *Fermi*-LAT detection of this pulsar wind nebula and the source spectrum appears to be consistent with a standard power law. An upper limit calculation predicts only about 100 events with energy above 1 GeV in the 9-year data set so a pulsation search was not conducted.

Introduction

The Large Area Telescope (LAT) is a gamma-ray detector, one of the instruments on the *Fermi* satellite, specifically sensitive to photons in the energy range ~ 10 MeV to > 300 GeV where eV is electronvolt. Gamma rays that enter the detector are converted by tungsten foil to positron-electron pairs which are then tracked by silicon strip detectors. The telescope was launched in 2009 and by mid 2017 had detected over 200 million photon events. It orbits the earth every 96 minutes with the entire sky being surveyed in two orbits. The goal of this paper is to investigate very-high energy (VHE) astrophysical sources using publicly available *Fermi*-LAT data and Science Tools[1]. Pulsars are one example of an astrophysical source which can emit in the VHE range. Pulsars are highly magnetized rotating stars (specifically neutron stars or white dwarfs) which emit beams of electromagnetic radiation which can be observed when the beam is pointing towards Earth, which happens every rotational period, hence the pulsed nature and the name of the stars. Pulsars have a magnetosphere, which is the region of space surrounding a source in which charged particles are controlled by the sources magnetic field.(6) A pulsar wind nebula is a region of particles accelerated in a shock formed by the interaction between the pulsar's particle or magnetic flux and surrounding matter.(1)

PSR J1930+1852 is young, energetic non third *Fermi*-LAT point source catalog (3FGL) (2)[2] pulsar. The first evidence of the pulsar was seen in 2002 in radio and x-ray emissions,(7) but its associated pulsar wind nebula G54.0+0.3 was observed in 1985.(8) It has a characteristic age (approximate age based on the slowing of the rotation of the pulsar) of 2900 years. (7) Both of these anomalous attributes are shared with the Crab pulsar, which has a characteristic age of 1257 years and emits in the VHE range. (9) The Crab pulsar also has a rare broken power law spectral energy distribution (SED) as seen by VERITAS (Very Energetic Radiation Imaging Telescope Array System, a ground-based gamma-ray observatory) in 2011.(10) Because of the other properties the Crab and PSR J1930+1852 share, the possibility of a broken power law SED in PSR J1930+1852 was investigated.

A source SED is the change of the flux with increasing energy, and a broken power law is one possible mathematical description of this relationship. In contrast to a standard power law in which the flux is proportional to the energy by some power or index, a broken power law has some cut off energy at which the SED has a different index, as illustrated below (Fig. 1).

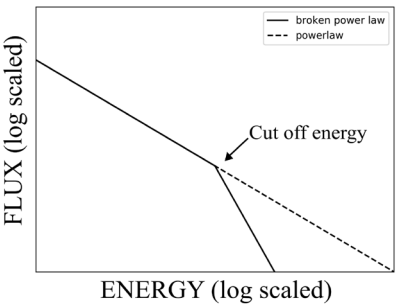


Fig. 1. An illustration of a spectrum in units of flux or plot of flux versus energy to illustrate the difference between a power law and a broken power law.

The detection of another pulsar with a broken power law in its VHE SED could add evidence for the mechanism for which pulsars are able to emit in the gamma-ray band. For example, there are models in which inverse Compton upscattering is the main emission mechanism which would be strongly supported by the detection of pulsars with broken power law spectra.(5) The analysis of G54.1+0.3 SED is also interesting because current measurements by different instruments operating at higher energies are in disagreement;(2,3) one of the goals was to better understand the SED behaviour across a broad energy spectrum. It has been seen by oth-

1 Data and Science Tools are available at https://fermi.gsfc.nasa.gov/ssc/data/access/.
2 The Third LAT Point Source Catalog, often abbreviated to 3FGL, is a catalog of over 3000 sources released by the LAT team after approximately 4 years of data taking.

er terrestrial detectors operating at higher energies, VERITAS (3) and HAWC (4) (High Altitude Water Cherenkov Observatory, a gamma-ray and cosmic ray observatory), but it was previously undetected in 9 years of *Fermi*-LAT data.

Methods

All *Fermi*-LAT photons from a 2.5-degree radius circular region near PSR J1930+1852 (specifically centred at our first estimation of its location based at right ascension 292.62 and declination +18.87) were selected and a model was made of the expected contribution from each source in the region. The model included possible emission from PSR J1930+1852; comparison with data was done via maximum likelihood estimation. Maximum likelihood estimation is a method of estimating parameters of a given model based on a sample population, done by maximizing the likelihood of making the observations in the sample given the parameters[3]. Here, the likelihood is the log of the probability assuming Poisson statistics in each energy bin in the SED. *Fermi*'s binned maximum likelihood function was utilized in the maximum likelihood fitting.

The steps of the fitting were to first make a data subselection around the region of interest and cut the data based on event class to ensure quality of the event data and zenith angle to discard non-astrophysical photons. The photon event files are from *Fermi* data version pass 8 between 239557417 s mission elapsed time (MET) and 518061255s MET. Next, one of the source model files (detailed below) was created from sources in the data region and up to 10° around it. Then the diffuse source responses were computed which depend on the instrument response function and finally the maximum likelihood fit was run. This final step was done twice, first to estimate the free parameters using a faster but less accurate optimizer (DRMNFB) and, from these estimates, apply a more accurate and CPU intensive optimizer (NEWMINUIT). DRMNFB finds the local minima of a continuously differentiable function subject to simple upper and lower bound constrains. NEWMINUIT is an interface to the C++ version of MINUIT, a well-known physics analysis tool for function minimization from CERN. This process was performed 3 times, once for each model. Test statistic maps, which search for unmodelled sources by calculating the likelihood of an additional source at each pixel of the map were made using another package in the Science Tools for the successful models. Finally, the upper limits of the flux from this source were computed assuming a power law model in Python (using *Fermi* Science Tools package). The other sources and their parameters were from 3FGL; all of the expected sources in the region and sources up to 10 degrees outside of the region were included, in case they contributed any photons as well. Three models were made; the first, shown in a counts map in the middle panel of the Fig. 2, is based on a model including only 3FGL sources, not the source of interest, PSR J1930+1852. The right panel of Fig. 2 is a model including all the 3FGL sources and a point source included for the pulsar with a fitted power-law spectrum. The equation for the power-law spectrum is

$$\frac{dN}{dE} = N_0\left(\frac{E}{E_0}\right)^{\gamma},$$

where N_0 is the prefactor, γ is the index, E_0 is the scale, and dN/dE is the integrated flux. The last model made was the same as the previously mentioned model except fitted with broken power-law source model. The equation for the broken power-law source model is

$$\frac{dN}{dE} = N_0 \times \begin{cases} (E/E_b)^{\gamma_1} & \text{if } E < E_b \\ (E/E_b)^{\gamma_2} & \text{otherwise} \end{cases},$$

where N_0 is the prefactor, E is the energy, E_b is the cut-off energy, dN/dE is the integrated flux, γ_1 is the index for energies less than the cut-off energy and γ_2 is the index for energies greater than the cut-off energy. The current *Fermi* galactic diffuse emission model and corresponding extragalactic isotropic diffuse emission were also accounted for and fit in all of these models. These background models and the 3FGL sources were fit with free amplitude parameters to account for possible variation. Then the maximum likelihood program was run on each of these models which would fit the background models, fit the source parameters and obtain their resulting "test statistics". The test statistic value is a way to quantify the quality of the maximum likelihood fit, and it roughly represents σ^2 significance for a normal ("Gaussian") distribution - so a larger value implies a higher likelihood of a gamma-ray signal. The formula for the test statistic is

$$TS = -2\ln\left(\frac{L_{max_0}}{L_{max_1}}\right),$$

where TS is the test statistic, L_{max0} is the maximum likelihood value for a model without an additional source and L_{max1} is the maximum likelihood value for a model with the additional source. Notice that the source of interest (just above and to the left of the center of the image) is far from being the brightest in the region.

Results

The precise results of the single power-law fit are as follows:

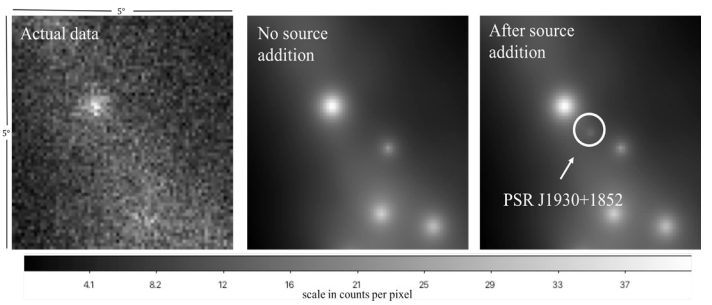


Fig. 2. Counts maps in the region of interest of (from left to right): *Fermi* data, the 3FGL point-source model without a new point source at the location of PSR J1930+1852 included, and the model with a source included (a power law pictured here, but a broken power law is visually the same). Each image is 5° × 5°.

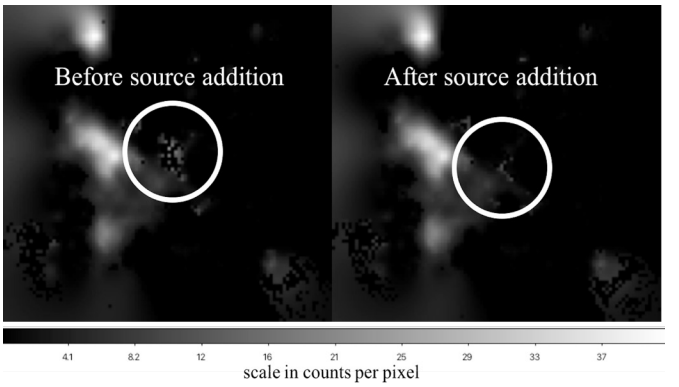


Fig. 3. Test Statistic map before and after adding to the model a source for PSR J1930+1852. Unexpected black pixels are likely due to the optimizer not converging in that pixel. If the model is an excellent fit to the data, one would expect the TS map to be close to featureless.

Prefactor (N_0): $0.03 \pm 0.03 \frac{\text{photons}}{cm^2 s}$

Index (γ): -1.9 ± 0.2

Scale (E_0): 100 MeV

TS value: 18.2 ($\sim 4.3\sigma$)

3 More information about maximum likelihood spectral fitting can be found on Fermi's website, https://fermi.gsfc.nasa.gov/ssc/data/analysis/documentation/Cicerone/Cicerone Likelihood/.

Note: The uncertainty on the prefactor appears to be consistent with zero. However, the test statistic was used to assert the source inclusion; the large uncertainty is a product of the log scale used, i.e. it reflects accurately that the prefactor could be as great as 0.06 photons/cm^{2}s. However, the range for values less than 0.03 photons/cm^{2}s is smaller.

Discussion

Source Inclusion Justification

Evidence of a new source (not in the 3FGL (1) catalog) can be seen in the test statistic maps (Fig. 3), which justifies the inclusion of a source. The test statistic maps are a measure of the probability that there is a source in the particular pixel that is not accounted for in the model, i.e. the expected contributions are already subtracted. In pixels with high significance where there was not a source modelled, this is an indicator that there may be an unmodelled source present. It can be seen near the region of interest that there is an excess of high significance in the fit without the added source, which - given the relatively low brightness of the source - is especially in excess. However, as shown in the second plot (Fig. 3), once the source corresponding to PSR J1930+1852 is added and fit with free parameters the excess is reduced, signifying a better fit. Any unexpected single black pixels in the Test Statistic maps (Fig. 3) are likely due to the optimizer not converging in that pixel. In making these maps the maximum likelihood program is essentially run in every pixel and since convergence is not guaranteed some number of these black pixels are to be expected.

Spectral Model Justification

Next, to conclude that the *Fermi* data supports a power-law fit for this source and not a broken-power law for the SED, there were two main arguments based on the TS value. When this region was fit using the broken power law model the cut off energy would remain at the lowest energy bin reducing the spectrum to a standard power law. When forced to pick a different cut-off energy by only allowing its energy cut-off parameter to be fit in a range above the lowest energy, the program would return a fit with a negative test statistic for the source of interest. Since the test statistic roughly corresponds to the significance squared a negative test statistic value is unphysical in nature and thus a sign of a bad fit. However, when this region is fit with a power law for the SED, the optimizer would converge and a test statistic corresponding to the 4.3 σ observation was returned. The index of the successful model was -1.9 ± 0.2, which is consistent with VERITAS reported index of SNR G54.1 + 0.3, $-2.39 \pm 0.30_{sys}$ from 100 GeV to above 30 TeV (3) but not consistent with HAWCs PSR J1930 + 1852 observation of an index of -2.74 ± 0.12 above 1 TeV.(4)

Upper Limit / Pulsed Signal

As this source is believed to be a pulsar, the next desired analysis was a pulsed signal search. A pulsed analysis was not done, however, because the upper limit flux calculation assuming the power-law model predicts the detection of only about 100 events from this source above 1 GeV in the entire 9-year *Fermi*-LAT data set. Although radio ephemerides[§] exist for PSR J1930+1852, this number of events in the nine-year data set was too small to detect, in a statistically significant way, pulsed emission. However, this is an investigation that is worthwhile for future analysis. Without searching for a pulsed signal it cannot be stated whether the emission is from the pulsar or the associated pulsar wind nebula.

Conclusion

Strong evidence is presented for a *Fermi*-LAT detection of PSR J1930+1852 and its associated pulsar wind nebula GN54.1+0.3, previously unseen in *Fermi*'s data set. This source spectrum appears to be consistent with a single power law rather than the broken power law originally suspected. This contributes to the knowledge of the source SED and may help to resolve the current uncertainties. A limitation to this research is that all sources modelled were considered as point sources. Although the high significance suggests that this is a good approximation, for completeness and further analysis, models including 3FGL extended sources in the region would be accounted for.

Another obvious extension of this research would be to search for a pulsed signal. However, given that an upper limit of the flux predicts approximately 100 events expected in the full 9-year data set, if pulsations were detected in this data after folding the data using an ephemeris derived from radio observations that is valid for the duration of the Fermi mission, it would likely be weakly. This can be considered as one of the drawbacks to the all-sky views that *Fermi*-LAT produces. If further research were to be conducted on this source, it would be recommended to use a more sensitive instrument with a deeper exposure, search this data for a pulsed signal, and use this combined with a pulsed radio detection to derive information about the magnetosphere of the pulsar.

Acknowledgements

I would like to thank my supervisors Prof. Ken Ragan and Dr. Ben Zitzer for their support, inspiration and mentorship.

I would also like to thank the NSERC URSA program for providing the funding for this research, and the McGill Physics Department for providing a nurturing environment.

References

1. H. An, K. Madsen, S. Reynolds, V. Kaspi, F. Harrison et. al. High-Energy X-Ray Imaging of the Pulsar Wind Nebula MSH 15-52: Constraints on Particle Acceleration and Transport. Astrophys. J. 2014 Oct 1; 793(1):90
2. F. Acero, M. Ackermann, M. Ajello, A. Albert, W. B. Atwood et. al. Fermi Large Area Telescope Third Source Catalog. Astrophys. J. Suppl. 2015 Jun 22; 218(2):23
3. V. A. Acciari, E. Aliu, T. Arlen, T. Aune, M. Bautista et. al. Discovery of Very High Energy γ-ray Emission from the SNR G54.1+0.3. Astrophys. J. Let. 2010 Jul 22; 719(1):L69
4. A. U. Abeysekara, A. Albert, R. Alfaro, C. Alvarez, J. D. A Ivarez et. al. The 2HWC HAWC Observatory Gamma-Ray Catalog. Astrophys. J. 2017 Jun 29; 843(1):40
5. Lyutikov, M. and Otte, N. and McCann, A. et. al. The Very High Energy Emission from Pulsars: A Case for Inverse Compton Scattering. Astrophys. J. 2012 Jul; 754(1):30
6. Anatoly Spitkovsky Pulsar Magnetosphere: The Incredible Machine. AIP Conference Proceedings 2008 Jul; 983(1):20-28
7. F. Camilo, D. R. Lorimer, N. D. R. Bhat, E. V. Gotthelf, J. P. Halpern, et. al. Discovery of a 136 Millisecond Radio and X-Ray Pulsar in Supernova Remnant G54.1+0.3. Astrophys. J. 2002 Jul 20; (1)71-74
8. Reich, W., Fuerst, E., Altenhoff, W.J., Reich, P., Junkes, N, et. al. Evidence for two young galactic supernova remnants. Astronomy and Astrophysics 1985 Oct; 151(1)10-12
9. Na Wang et. al. A Large Glitch in the Crab Pulsar. Chin. J. Astron. Astrophys. 2001 Apr; 1(3)95–199
10. A. McCann. Detection of the Crab Pulsar with VERITAS above 100 GeV. ICRC Proceedings 2011 Oct 19; 32(1)

Submitted: 10/03/17

Research Article

[1]Department of Psychology, McGill University, Montreal, QC, Canada

Keywords

Blood donation, Vasovagal Syncope, Respiration Control, Applied Muscle Tension

Email Correspondence

kristina.wade@mail.mcgill.ca

Kristina Wade[1], Blaine Ditto[1]

Effectiveness of Behavioural Intervention as Treatment for the Vasovagal Response in Blood Donation

Abstract

Background: The experience of a vasovagal reaction during blood donation, with symptoms such as dizziness, weakness, and fainting, contributes to a more negative donation experience and significantly decreases the likelihood of blood donor return. This study investigates the effects of two behavioural interventions on reducing the occurrence of such reactions, applied muscle tension and respiration control, and possible moderation of these effects by sex, BMI, and medical fear.

Methods: Six hundred and eleven participants were recruited from Héma-Québec blood drives across Montreal and randomly assigned one of four conditions: applied muscle tension, an anti-hyperventilation respiration control procedure, both techniques, or neither. Following their donation, participants completed the Blood Donations Reactions Inventory and Medical Fears Survey. Analysis focuses on the respiration control and applied tension groups.

Results: While donor sex and BMI did not predict the effectiveness of applied muscle tension intervention, results showed that the largest benefit was seen in donors who reported lower levels of medical fears in the respiration control condition group.

Limitations/Conclusions: The results are promising in that they suggest that intervention can decrease the risk for vasovagal symptoms in blood donation, though it may not be sufficient to reduce symptoms in donors with high levels of medical fear.

Introduction

The need for a large and consistent blood supply is critical for the medical system to function effectively. Health care relies heavily on transfusions of whole blood and its components to complete many types of necessary medical interventions.(1) Maintenance of an adequate blood supply at all times can only be achieved through sufficient numbers of voluntary donations from healthy individuals. Even though about 60% of the population of North America is eligible to donate blood, an analysis revealed that only five percent of the population donates frequently.(2) Increasingly rigid criteria for donation also limits the pool of potential donors. In 1988, a total of 13 exclusion criteria left 64% of the U.S. population eligible to donate blood.(3) This Fig. dropped to approximately 38% in 2007, when the total number of exclusion criteria reached 31.(4) Clearly, given the demands of our aging population, the limited pool of donors who meet criteria, and the limited number of potential donors who choose to donate, the pressure to increase blood donation is evident. A greater emphasis needs to be placed on recruiting new donors, as well as encouraging repeat donation, in order to maintain adequate blood reserves. One of the biggest obstacles to blood donor retention is the experience of vasovagal symptoms such as fainting, dizziness, weakness and nausea.

Mechanisms of vasovagal syncope

Research suggests that vasovagal symptoms may occur in response to an exaggerated stress reaction.(5) Following the experience of an anxiety-inducing stimulus such as blood donation, the sympathetic nervous system is activated, which increases heart rate and blood pressure. However, in periods of prolonged and intense anxiety, the body cannot maintain this level of arousal. As a result, the parasympathetic nervous system is triggered, leading to a rapid decrease in heart rate and blood pressure, which can be exaggerated at times. There are other possible mechanisms of the vasovagal response, but this a common explanation.(6) This sharp change in blood pressure and heart rate can result in loss of consciousness, weakness, and lightheadedness. Studies have shown that adverse reactions such as these make blood donors less willing to donate again in the future.(7,8)

Treatment

Methods of preventing these reactions may increase donor retention. The applied muscle tension (AT) technique has been shown to reduce vasovagal reactions in many studies.(9) This technique was originally developed as a behaviour therapy intervention to treat blood and injury phobias, but it quickly became apparent that muscle-tensing techniques could be useful for non-phobics as well, as in the case of blood donors.(9) The AT technique consists of repeated five second cycles of isometric muscle tension of the major muscle groups in the arms and legs.(10) It is presumed that this technique reduces vasovagal symptoms by increasing blood pressure and cerebral blood flow.(11)

It is hypothesized that the effects of the AT technique may be mediated in part by encouraging regular breathing. When people learn AT, they are usually told to maintain regular breathing patterns during muscle tension. Among other possible mechanisms, it has been suggested that vasovagal reactions may be exacerbated by hyperventilation that can be induced by stressful or anxiety-producing situations. Hyperventilation decreases the amount of carbon dioxide in the blood and if these levels become low enough, the body can enter a state of hypocapnia, which causes the blood vessels in the brain to constrict.(12) This vasoconstriction may cause a decrease in cerebral blood flow, reducing the amount of oxygen delivered to the brain. Thus, respiration control techniques have been developed to reduce the occurrence of hypocapnia episodes, and increase cerebral blood flow.(13) Although the effects of such techniques have not been studied in blood donors, they have been implemented in patients with blood-injection phobia. A recent study by Ritz suggested that a shallow breathing technique to prevent hyperventilation was as effective as an AT technique in reducing vasovagal symptoms in a sample of blood-injection phobics. (13) We intend to investigate the efficacy of both of these treatments in the context of blood donation.

Potential Moderators

While previous studies have shown that applied muscular tension and respiration control techniques can be effective at reducing symptoms of dizziness, weakness and fainting, various traits have been shown to affect the likelihood of experiencing an adverse reaction, as well as influence a patient's response to treatment.(9,13)

Donor Sex

In general, women are more likely than men to report symptoms during blood donation.(14) Additionally, in a study by Kamel, it was shown that delayed reactions are three times more likely to occur in females than males.(15) Thus, female sex seems to be a risk factor for vasovagal reactions during blood donation. This association might be explained in part by the fact that women tend to have a lower body mass than men, and a lower body mass is usually associated with a lower estimated blood volume (EBV). By calculating EBV using self-reported weight, height and sex, studies have shown that donors with lower EBV's have an increased risk of vasovagal reactions.(15) Because all donors are required to donate the same amount of blood, this might lead to increased vasovagal reactions in female participants. In line with this reasoning, Ditto found that AT significantly reduced symptoms in women but not men.(11)

Body Mass Index (BMI)

Newman found that low weight (<130 pounds) was associated with a vasovagal reaction rate of 13.6%, compared to 3.3% in higher weight donors.(16) Similar to EBV, BMI can be calculated using self-reported measures of weight, height and sex of blood donors. While national health recommendations indicate that BMI's of 18.5-24.9 are within the healthy range, blood donors with a BMI < 25 have been shown to be at greater risk for experiencing vasovagal reactions.(17) Low BMI is associated with low blood pressure, which in turn is related to susceptibility to vasovagal reactions. Blood donation causes blood pressure levels to drop; individuals who have low blood pressure at baseline might be more at risk to experience symptoms of dizziness, weakness, or fainting following donation. Given that applied muscle tension has the capacity to maintain blood pressure levels, this technique might be especially useful in donors with lower BMI's. Because of this, we will investigate whether BMI moderates the effects of applied muscle tension.

Medical Fears

In addition to biological risk factors for vasovagal reaction, donors' levels of fear may influence their likelihood of experiencing a vasovagal reaction. In fact, fear of injection and blood has been shown to be the strongest predictor of vasovagal symptoms in women, while the same was true for men who were first time donors.(18) As previously discussed, an anti-hyperventilation respiration control technique has been shown to be effective in reducing vasovagal symptoms in a sample of blood-injection phobics. Such a technique might be predicted to work better with high fear donors, if it successfully controlled respiration rates. Further, research concerning the efficacy of similar treatment in a sample of blood donors might be valuable in terms of targeting treatment interventions to higher risk groups, in order to reduce the probability of experiencing a reaction, thus increasing the likelihood of repeat donation.

The Current Study

This study was part of a randomized control trial investigating the effectiveness of three interventions in reducing the occurrence of vasovagal reactions during blood donation; applied muscular tension, respiration control, and a combined technique. Analyses will focus on results from the applied muscle tension and respiration control groups. It was hypothesized that applied muscle tension and respiration control would be effective at controlling vasovagal symptoms and reducing the need for treatment. Further, effects of applied muscle tension would be moderated by sex and BMI, whereas respiration control would be moderated by preexisting levels of medical fears.

Methods

Participants

Participants were recruited from mobile Héma-Québec blood drives at various universities and CEGEPs (collège d'enseignement général et professionnel; students aged 18 to 20 years) across Montreal, Québec. A total of 611 donors (321 female and 281 male) between the ages of 18 and 35 were recruited (Mean = 21.7, Standard Deviation = 3.3 years). Participants were assigned randomly to one of four treatment conditions; an applied tension only group (151 donors), a respiration control only group (153 donors), an applied tension and respiration control group (153 donors), and a control group (154 donors).

Procedure

After obtaining informed consent, participants completed a predonation questionnaire requesting demographic information and mood ratings. They were then randomly assigned one of four conditions. Based on the condition they were assigned, participants were given padded earphones to watch an instructional video. Each video demonstrated the technique that participants were asked to practice before and during blood donation. The videos were available in both English and French, narrated by the same bilingual narrator. The first video corresponded to the Applied Muscle Tension condition, during which participants were shown a muscle tensing technique. In the Respiration Control condition, participants were shown a shallow, anti-hyperventilation breathing technique. The third intervention was a combination of the Applied Muscle Tension technique and the Respiration Control technique. Participants in the Control condition did not watch an instructional video, but proceeded to the next set of measurements. After the video, a research assistant measured each participants' blood pressure and heart rate. Next, a portable capnometer (Micrograph Plus, Oridion Capnography, Minneapolis, MN) was attached to each participant to measure end-tidal CO_2, an index of hyperventilation, throughout the procedure. Participants then continued with the standard donation process.

A research assistant was present in the donation area to observe the behaviour of the participants, and to verify that they practiced the assigned technique(s) during the donation process. The research assistant completed an assessment form, indicating whether a vasovagal reaction occurred, whether the reaction required treatment, and any other irregularities witnessed throughout the procedure. When the donation was complete, participants were asked to wait in a rest area for five minutes. Following this rest period, the capnometer was removed, and participants proceeded towards a post donation snack area, where they were asked to complete a post donation questionnaire. This questionnaire included the Blood Donations Reactions Inventory (BDRI) to evaluate participants' experience of vasovagal symptoms, assessed the participants' likelihood of donating again, asked the degree to which the participant practiced their assigned technique (all the time or occasionally before, during or after donation), and included the Medical Fears Survey.(18,19) Finally, a research assistant obtained post donation measures of blood pressure and heart rate.

Experimental Conditions

Applied Muscle Tension

The Applied Muscle Tension video demonstrated a whole body, isometric muscle tension technique. Participants were asked to engage in repeated cycles of tension for five seconds, followed by five seconds of rest. They were told to emphasize leg muscle tension by pointing their toes downwards and flexing their thigh muscles. Participants were reminded to breathe normally throughout their donation.

Respiration Control

Individuals in this condition were shown an anti-hyperventilation breathing technique, modeled by the narrator of the video. Participants were asked to engage in slow, shallow breathing with their mouths closed. They were also instructed to use abdominal breathing, rather than chest breathing. This technique was emphasized by asking the participants to place one hand over their chest, and the other over their abdomen, allowing them to

observe the movement of their diaphragm while breathing.

Combined Technique

In this video, the Applied Muscle Tension technique was explained first, followed by the Respiration Control technique. Participants were asked to practice both techniques during their donation.

Measures

Blood Pressure and Heart Rate

Pre and post donation measures of blood pressure and heart rate were obtained using manual blood pressure monitors (Model A10, Becton Dickinson, Franklin Lakes, NJ).

Capnometer

Participants were asked to wear a portable capnometer (Micrograph Plus, Oridion Capnography, Minneapolis, MN) throughout the study in order to obtain end- tidal CO_2 measurements. Capnometer data was not analyzed for the purposes of the current study.

BDRI

The Blood Donation Reactions Inventory (BDRI) is a well-validated, 11-item scale that allows participants to self-report the experience of vasovagal symptoms during donation on a 6-point scale. A shortened four-item version of the BDRI has shown to be useful in assessing the subjective perception of pre-faint symptoms; therefore, this version was used for statistical analysis.(19) This version relies on the results from items 1, 2, 3, and 7, looking at participant self-report of faintness, dizziness, weakness, and lightheadedness, respectively.

Medical Fears Survey

This self-reported survey contains 25 items. In various ways, participants were asked to indicate the level of fear or tension they would likely experience when exposed to different medical situations, such as having blood drawn from the arm.(20)

Results

A total of 521/611 individuals completed their blood donations and had data entered in the analyses. Participants who did not meet the eligibility requirements issued by Héma-Québec were not able to complete their blood donation, and therefore had to be excluded from the study.

Donor Sex

An analysis of variance revealed that women were significantly more likely to report symptoms on the BDRI than men, $F(1,585)=13.928$, $p=0.000$, however, women were not more likely to require treatment during donation than men. Results indicated that the effects of AT were not moderated by sex. A Pearson correlation matrix indicated that donor sex and scores on the Medical Fears Survey were moderately negatively correlated, such that men were slightly less likely to report medical fears, $r(581)=-0.354$, $p=0.000$.

BMI

Donors' BMI was calculated using self-reported weight, height, gender and published formulas. An analysis of variance revealed that donors with low BMI were significantly more likely to report symptoms on the BDRI, $F(1,591)=10.229$, $p=0.001$, although they were not more likely to require treatment by a nurse. Further, the effects of AT were not moderated by BMI.

Medical Fears

For this measure, the primary analyses were 2 Respiration Control (yes/no) x 2 Applied Tension (yes/no) x Total Medical Fears Score (treated as a continuous variable) general linear models (GLMs). The primary dependent variables were BDRI score and whether or not the donor was treated by the nurse for a vasovagal reaction as indicated by the research assistant. Both analyses produced strong main effects of Medical Fears, $F(1,490)=39.90$, $p<0.001$ and $F(1,484)=14.50$, $p<0.001$, respectively. In general, donors who reported greater fears in medical environments were much more likely to indicate symptoms of dizziness, weakness, lightheadedness and faintness on the BDRI (Fig. 1) and more likely to receive treatment for a vasovagal reaction (Fig. 2).

Figure 1: BDRI Score Outcomes by Medical Fear

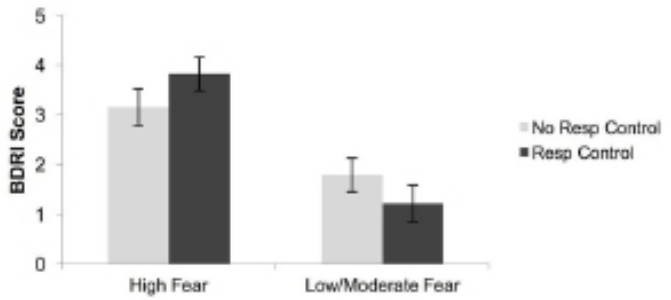


Fig. 1. Donors reporting greater fears in medical environments were more likely to indicte symptoms of dizziness, weakness, light-headness and faintness on the BDRI than donors reporting lesser fears.

Figure 2: Treatment for Vasovagal Reactions by Medical Fear

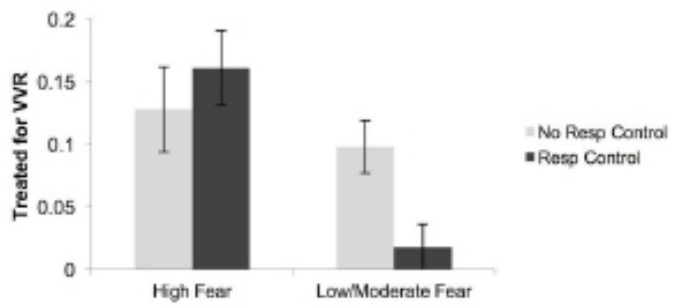


Fig. 2. Donors reporting greater fears in medical environment were more likely to receive treatment for vasovagal reactions than donors reporting lesser fears.

The effects of medical fears were influenced to some degree by the respiration control intervention as indicated by significant 2-way Medical Fears x Respiration Control interactions in both analyses, $F(1,490)=9.36$, $p=0.002$ and $F(1,484)=11.30$, $p=0.001$, respectively. There were no significant effects involving applied tension in the analyses, and this intervention will not be discussed further.

As can be seen in Fig. 1, the BDRI interaction was due to a combination of somewhat fewer (but not significantly fewer) symptoms in lower fear participants who practiced respiration control and somewhat more (but not significantly more) symptoms in higher fear participants who practiced respiration control.

As can be seen in Fig. 2, while the same non-significant trend was observed among high fear donors in regards to whether or not they required treatment for a vasovagal reaction, the benefit of practicing respiration control was much stronger among donors who were somewhat less fearful. In fact, after dividing participants based on the median medical fear score, less fearful donors who practiced respiration control were significantly less likely to require treatment for a vasovagal reaction compared to those who did not practice the technique, $F(1,244)=7.15$, $p=0.008$.

Discussion

The primary goal of this study was to investigate the effectiveness of the two chosen behavioural interventions as possible treatments for vasovagal reactions during blood donation. It was hypothesized that findings from previous studies would be replicated, indicating that females would be more likely to experience vasovagal symptoms than males. Expectations were that donor sex and BMI would moderate the effectiveness of applied muscle tension, while medical fears would moderate the effectiveness of respiration control.

Results were consistent with findings from previous studies, indicating that being female would present as a risk factor for the occurrence of a vasovagal reaction.(14) While female donors reported higher scores on the BDRI, they were not significantly more likely to be treated than men. This might be because while women were more likely to experience symptoms in the clinic than men, this difference was not pronounced enough for women to require more treatment. Alternatively, women may simply be more likely to report symptoms retrospectively, whereas men may be less likely to report them. Thus, a bias in self-report of symptoms could influence results on the BDRI. These factors may also explain why the effects of AT were not moderated by sex. If, in absence of treatment conditions, men and women do not differ in likelihood to require treatment during blood donation, then applied tension is unlikely to benefit one sex more than the other in terms of reducing the occurrence of symptoms.

Further, donors with low BMI reported experiencing more symptoms on the BDRI questionnaire than high BMI donors. BMI did not influence the likelihood to require treatment, nor did it moderate the effects of AT.

Another aim of this study was to investigate the benefits of implementing an anti-hyperventilation respiration control technique to reduce vasovagal reactions in blood donation. As demonstrated by results on the BDRI as well as observational measures made by a research assistant during donation, it was expected that vasovagal reactions would be more likely to be experienced by donors with higher levels of medical fears. The effects of the respiration control technique were hypothesized to be particularly helpful for this group of donors because they would be the most likely to hyperventilate.

Results support the hypothesis that levels of medical fears influenced the occurrence of vasovagal reactions. Donors who reported high levels of fear in medical situations were significantly more likely to report experiencing vasovagal symptoms during blood donation. These donors were also significantly more likely to be treated for a reaction during their donation.

Results failed to support the hypothesis that the benefits of respiration control would be highest amongst donors with high levels of medical fear. Instead, outcomes of the BDRI indicate that amongst donors in the respiration control treatment group, lower fear participants reported slightly fewer vasovagal symptoms, whereas higher fear participants reported slightly more symptoms. It is possible that presenting this intervention to high fear donors heightened their fears by drawing attention to the possibility of an adverse reaction. Alternatively, perhaps they found the task too complicated, which reduced the positive effects of the treatment. Similarly, participants with lower fear levels were less likely to require treatment for a vasovagal reaction if they were in the respiration control group. These results suggest that recommendations to practice the anti-hyperventilation breathing technique should be aimed at donors who initially report low levels of fear or tension with regards to medical situations.

The correlation between donor sex and Medical Fears Survey scores was calculated to investigate whether medical fears in the respiration control group were influenced by donor sex. Males were slightly less likely to report medical fears. While this was a subjective measure, given that analysis revealed only moderate correlations, future studies should look into this further.

A few limitations exist with regards to the research and analysis of this study. Firstly, while the sample size was large, it was mostly comprised of students due to convenience sampling. Participants were recruited from mobile Héma-Québec blood drives at CEGEPs and universities across Montreal, making student participation predominant. This limits the capacity to generalize our findings to older donor populations.

The strict eligibility criteria of Héma-Québec resulted in 90 participants being deferred from blood donation, making them unable to complete the study and be included in analysis. This reduction in sample size may have prevented strong correlations from being seen in the data.

A further limitation of this study was that both the Medical Fears Survey and the BDRI questionnaire were completed post-donation. As such, donors' self-report of medical fears might have been influenced by their experience during blood donation. Being exposed to some of the aspects discussed in the survey during the course of the donation, such as needles and blood, may have temporarily heightened individuals' levels of fear. Filling out the surveys after the experience of a vasovagal reaction may have also made participants more likely to report higher levels of medical fears. Subsequent research should investigate whether the administration of the Medical Fear Survey prior to donation influences results, although exposing donors to the information on this survey might end up priming fear or anxiety by suggestion. An interesting alternative would be to administer the survey both prior to and following donation, to compare scores and see whether they remain consistent.

Similarly, the BDRI requires participants to report symptoms such as dizziness, weakness, and faintness on a 6-point scale. The completion of this survey after the donation process may have caused participants to be less accurate in reporting the extent to which they experienced symptoms. While results might be more accurate if the BDRI could be completed in real time, this is highly unrealistic given the complexity of the donation process in the clinic. Participants would most likely not be able to complete the survey on their own, and if a research assistant asked the donor direct questions for an extended period of time, it may interfere with the work of the nurses and other clinic personnel.

Conclusion

This study was part of a randomized control trial investigating the effectiveness of three interventions – applied muscular tension, respiration control, and a combined technique – as treatment for the vasovagal response in blood donation. The goals were to assess the effectiveness of applied muscular tension and respiration control on reducing vasovagal reactions and treatment likelihood, as well as to investigate the moderating effects of donor sex and BMI on applied muscle tension and the effect of medical fears on respiration control. While donor sex and BMI did not moderate the effects of AT, there was a significant interaction between medical fears and respiration control. It seems as though respiration control might be most useful in reducing vasovagal symptoms of donors with low medical fears. Given these findings, blood collection agencies might suggest that donors self-reporting low medical fears practice the respiration control technique during the donation process in order to reduce the likelihood of experiencing a reaction. In doing so, this intervention has the potential to increase the occurrence of repeat donation, and contribute to the maintenance of adequate blood reserves for our health care system.

Acknowledgements

The author thanks Dr. Blaine Ditto for his guidance and expertise, and for aiding in data analysis. The author also thanks Dr. David Ostry for providing commentary that helped improve the manuscript, and the nurses and staff of Héma-Québec for their assistance with this research.

References

1. Sullivan, M., Cotten, R., Read, E., Wallace, E. Blood Collection and transfusion in the United States in 2001. Transfusion. 2007;47(3):385-394.
2. Wu, Y., Glynn, S., Schreiber, G., Wright, D., Lo, A., et al. First time blood donors: demographics trends. Tranfusion. 2001;41(3):360-364.
3. Linden, J., Gregoria, D., Kalish, R. An Estimate of Blood Donor Eligibility in the General Population. Vox Sang. 1988;54(2):96-100.

4. Riley, W., Schwei, M., McCullough, J. The United States' potential blood donor pool: estimating the prevalence of donor-exclusion factors on the pool of potential donors. Transfusion. 2007;47(7):1180-1188.

5. Graham, D. Prediction of fainting in blood donors. Circulation. 1961;23:901-906.

6. Graham, D.T., Kabler, J.D., Lunsford, L. Vasovagal Fainting: A Diphasic Response. Psychosom Med. 1961;23(6):493-507.

7. Olatunji, B.O., Etzel, E.N., Ciesielski, B.G. Vasovagal syncope and blood donor return: examination of the role of experience and affective expectancies. Behav Modif. 2010;34(2):164-174.

8. France, C.R., Rader, A., Carlson, B. Donors who react may not come back: analysis of repeat donation as a function of phlebotomist ratings of vasovagal reactions. Transfus Apher Sci. 2005;33(2):99-106.

9. Ditto, B., France, C.R., Albert, M., Byrne, N. Dismantling applied tension: mechanisms of a treatment to reduce blood donation-related symptoms. Transfusion. 2007;47(12):2217-2222.

10. Holly, C.D., Balegh, S., Ditto, B. Applied tension and blood donation symptoms: the importance of anxiety reduction. Health Psychol. 2011;30(3):320-325.

11. Ditto, B., France, C.R., Lavoie, P., Roussos, M., Adler, P.S. Reducing reactions to blood donation with applied muscle tension: a randomized controlled trial. Transfusion. 2003;43(9):1269-1275.

12. Alboni, P., Brignole, M., Uberti, E.C. Is vasovagal syncope a disease? Eurospace. 2007;9(2):83-87.

13. Ritz, T., Meuret, A.E., Ayala, E.S. The psychophysiology of blood-injection phobia: Looking beyond the diphasic response paradigm. Int J Psychophysiol. 2010;78(1):50-67.

14. Byrne, N., Ditto, B. Alexithymia, cardiovascular reactivity, and symptom reporting during blood donation. Psychosom Med. 2005;67(3):471-475

15. Kamel, H., Tomasulo, P., Bravo, M., Wiltbank, T., Cusick, R., et al. Delayed adverse reactions to blood donation. Transfusion. 2010;50(3):556-565.

16. Newman, B.H. Vasovagal reactions in high school students: findings relative to race, risk factor synergism, female sex, and non-high school participants. Transfusion. 2002;42(12):1557-1560.

17. Takanshi, M., Odajima, T., Aota, S., Sudoh, M., Yamaga, Y., et al. Risk factor analysis of vasovagal reaction from blood donation. Transfus Apher Sci. 2012;47(3):319-325.

18. Labus, J.S., France, C.R., Taylor, B.K. Vasovagal reactions in volunteer blood donors: analyzing the predictive power of the medical fears survey. Int J Beh Med. 2000;7(1):62-72.

19. France, C.R., Ditto, B., France, J.L., & Himawan, L.K. Psychometric properties of the Blood Donation Reactions Inventory: a subjective measure of presyncopal reactions to blood donation. Transfusion. 2008;48(9):1820-1826.

20. Olatunji, B., Ebesutani, C., Sawchuk, CN., McKay, D., Lohr, J., et al. Development and initial validation of the medical-fear survey-short version. Assessment. 2012;19(3):318-336.

21. Labus, J.S., France, C.R., Taylor, B.K. Vasovagal reactions in volunteer blood donors: analyzing the predictive power of the medical fears survey. Int J Beh Med. 2000;7(1):62-72.

22. France, C.R., Ditto, B., France, J.L., & Himawan, L.K. Psychometric properties of the Blood Donation Reactions Inventory: a subjective measure of presyncopal reactions to blood donation. Transfusion. 2008;48(9):1820-1826.

23. Olatunji, B., Ebesutani, C., Sawchuk, CN., McKay, D., Lohr, J., et al. Development and initial validation of the medical-fear survey-short version. Assessment. 2012;19(3):318-336.

Submitted: 12/25/17

Research Article

[1]Department of Chemistry, McGill University, Montreal, QC, Canada

Keywords

Superparamagnetic, nanoparticle, ligand

Email Correspondence

hannah.sragovicz@mail.mcgill.ca

Hannah Sragovicz[1]

Techniques for Surface Modification of Aqueous-Stable Superparamagnetic Iron Oxide Nanoparticles

Abstract

Background: The iron oxide nanoparticles involved in this study are unique in their superparamagnetic properties, defined as their ability to flip the direction of their magnetic field under influence of temperature. This property has a variety of environmental and biomedical uses. Indeed, the exchange of ligands on the surface of these particles enables exploration of such applications. The purpose of this study is to determine an efficient method of ligand exchange in order to standardize the surface modification of these iron oxide nanoparticles (IONPs). Namely, the primary methods of ligand exchange to be evaluated are shaking and sonication of reaction mixtures. As part of this method comparison, the exchange of oleic acid (OA) ligands for 3,4-Dihydroxyphenylacetic acid (DOPAC) ligands serves as a general model for method comparison. When comparing methods, both time and quantity of materials required are considered. The quality of the final product is also considered, assessed by factors such as oxidation state, colloidal stability, and extent of ligand exchange.

Methods: Three methods of ligand exchange are performed, after which their products are compared. The first method involves shaking the mixture overnight for a duration of 18 hr. The second method involves sonication for a duration of 30 min. The third method involves sonication of the reaction mixture for an additional 30 min. (duration of 60 min. in total).

Results: The products were analyzed using Fourier-transform infrared spectroscopy (FT-IR), zeta potential measurements, thermogravimetric analysis (TGA), and x-ray photoelectron spectroscopy. FT-IR measurements indicate that the one-time sonication method leads to the surface of the IONPs bearing the most residual oleic acid, a disadvantageous result. TGA analysis indicates that the twice-sonicated product is more favourable than the once-sonicated product.

Limitations: Larger data sets of FT-IR, TGA, zeta potential, and XPS must be collected before the best method may be confirmed. Zeta potential measurements must be repeated for the shaken product at a concentration that matches that of the other products. As such, a direct comparison may be made. TGA must also be repeated for the shaking product in order to eliminate possible inaccuracies. Namely, these could result from technical difficulties encountered in the measurement discussed above. While zeta potential measurements indicate that the twice-sonicated product has the highest colloidal stability, XPS measurements did not vary significantly enough between methods to suggest a most advantageous method.

Conclusion: According to the TGA and zeta potential measurements, the twice-sonicated product appears to be most favourable in terms of coverage. XPS suggests that all methods are comparable in terms of oxidation of the IONPs' iron.

Introduction

The use of aqueous stable superparamagnetic iron oxide nanoparticles (IONPs) is widespread. Biomedical and bioengineering applications include enhancement of magnetic resonance imaging contrast, tumor hyperthermia (1), drug delivery, tissue repair, and detoxification of biological fluids.(2) The potential for widely varying surface coatings of IONPs allows for widely varying applications. For example, IONPs may bind to proteins, antibodies, or drugs, as well as be redirected to specific tissues using an external magnetic field.(3) Other applications include the development of hybrid organic-inorganic materials.(4) Other applications outside biomedicine include data storage and water treatment.(5) Significantly, the preparation of aqueous-stable IONPs required by such applications generally involves some form of ligand exchange technique. The goal of this process is to optimize magnetic attributes of these nanoparticles, especially to improve hydrophilicity. In fact, this is particularly necessary for biomedical applications.

Such ligand exchange techniques include mechanochemical milling, shaking, and sonication. Mechanochemical milling is useful for circumventing issues such as solvent compatibility limitations, and also eliminates the need for chlorinated solvents.(6) Indeed, it is possible to employ mechanochemical milling to eliminate intermediate substitution steps in making superparamagnetic IONPs soluble in aqueous buffers. For instance, this allows for one-step conversion of monodisperse hydrophobic oleic-acid capped superparamagnetic IONPs to hydrophilic Tiron-capped IONPs. While useful, mechanochemical milling will not be considered in this particular comparison of ligand exchange methods, focusing rather on the shaking and sonication methods.

To this end, the ligand exchange of oleic acid (OA) ligands for 3,4-Dihydroxyphenylacetic acid (DOPAC) will serve as a general model for method comparison (Fig. 1).

This model proves to be especially useful because shaking and sonication methods are known to reliably produce IONP-DOPAC products. This exchange is favoured due to the catechol group's higher affinity for iron (III) as compared to that of the carboxylate.(7) The functional groups present

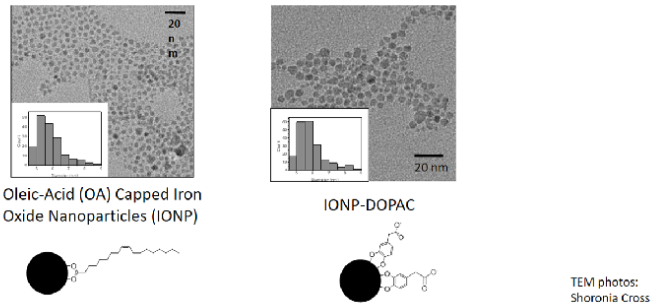


Fig. 1. TEM Photos of Starting Material and Product.

on the surface of the catechol impart aqueous stability. In other words, DOPAC binds through the catechol group, leaving the ionizable carboxyl group exposed to solution, thus making the particles aqueous-stable.

The first method evaluated, the shaking method, involves shaking the mixture overnight for approximately 18 hr.(8) On one hand, the shaking method features a significant benefit in its ability to carry out an almost complete ligand exchange. Nevertheless, a small amount of OA remains on the IONPs regardless of the method used, and the surface of the IONPs will always have OH groups present. Therefore, the surface will never be fully covered with any ligand.

On the other hand, this method disadvantages the process due to the amount of time it takes to reach completion. This could greatly encumber both the synthesis of IONPs functionalized with dopamine derivatives, and the subsequent tests to be performed on these products.

The two other methods to be examined involve sonication of the reaction mixture.(9) Previously it was theorized that the IONPs first undergo partial exchange, forming an intermediate that is covered partially by DOPAC and partially by OA.

In one sonication method, the mixture undergoes sonication for 30 min., which is favorable due to its short reaction time.(10) However, this technique can hinder the exchange process by its production of a compound which may not have optimal DOPAC coverage. Additionally, sonication's more energetic nature poses a greater risk of degradation or oxidation of the IONPs.

In the second method of this kind, the mixture is subjected to 30 min. of sonication, followed by a second addition of DOPAC to the particles, and another 30-min. period of sonication. A potential advantage of this method is an optimized addition of DOPAC to the surface of the IONPs. Still, this technique demands double the amount of ligand, and results in an inevitable loss of IONPs following the second sonication. Indeed, this purification (with washing) comes with a noticeable loss of IONPs. Furthermore, this method's second sonication risks causing more degradation and oxidation of IONPs.

The difference between theoretical products resulting from these methods is illustrated in Fig. 2. Both sonication methods are clearly favourable over shaking due to their comparatively short reaction time.

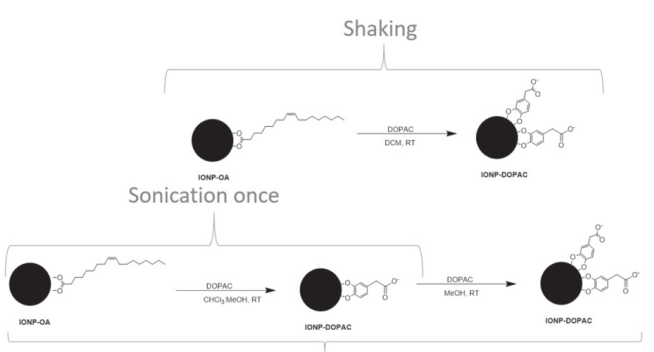


Fig. 2. Methods to be Compared.

Looking towards the future, the consistency resulting from a standardized methodology will enhance the reliability and usefulness of findings related to the various dopamine derivatives, which will be synthesized, in collaboration with the Lumb group, and analyzed. Afterwards, these dopamine derivatives will be used as a model for ligand electronic effects. As the aromatic ring will become more or less electron deficient depending on the ligand involved, the effect of this property on particles' binding behaviour and stability will be explored. For example, an electron-poor ligand may pull electrons off the surface iron, affecting its oxidation state and the crystallinity of the nanoparticle.

Methods

Ligand exchange via single sonication:(8) A room temperature 2 mL aliquot of a stock solution of 2.5 mg/mL IONP-OA in hexane was dried under nitrogen flow. The particles were then re-dispersed in 14 mL of chloroform by mild vortexing. 10 µL of DI water were added to 44.4 mg of DOPAC dissolved in 1 mL of methanol. This solution was added in 0.5 mL aliquots to the nanoparticle solution in chloroform. The mixture was then sonicated for 30 min. at room temperature.

Ligand exchange via double sonication: After the first 30 min. sonication, the mixture was pelleted by centrifugation at 4000 xg for 30 min. at 4°C. The supernatant was decanted, the pellet was re-dispersed in a solution of 44.4 mg of DOPAC in 10 mL of methanol, and the mixture was sonicated a second time for 30 min. at room temperature.

Ligand exchange via overnight shaking:(7) A 2 mL aliquot of a room temperature 2.5 mg/mL stock solution of IONP-OA in dry DCM was added to 44.4 mg of DOPAC dissolved in 13 mL of dry DCM. 10 µL of DI water were added to this reaction mixture, and the mixture was shaken for 18 hr at room temperature by vortex.

Purification of IONP-DOPAC products: All IONP-DOPAC products were purified identically, regardless of preparation. The reaction mixture was centrifuged at 4000 xg for 30 min. at 8°C. Supernatants were removed by pipet and pellets were re-dispersed in methanol. The mixture was again centrifuged at 4000 xg for 30 min. at 8°C. The supernatants were removed and the particles were dried under nitrogen flow. For FT-IR analysis, 20 µL of methanol was added to create a slurry for plating.

Fourier-Transform Infrared Spectroscopy (FT-IR): ATR-FTIR spectra were collected using a Spectrum Two FT-IR spectrometer equipped with a diamond ATR accessory and processed using Spectrum FT-IR software (PerkinElmer Inc. Waltham, MA, USA). All spectra were recorded between 4000 and 400 cm^{-1}, with 4 cm^{-1} resolution, averaged over 16 scans. A 1 mL aliquot of sample was centrifuged for 30 min. at 4400 rpm, supernatant was removed and methanol was added. The resulting mixture was plated directly onto the FT-IR as a film.

Zeta Potential: Before performing zeta potential measurements, it was necessary to determine the iron concentration via an established procedure, on an aliquot of each IONP solution in buffer, and to then dilute the sample to 0.05 mg/mL. Measurements were performed on a ZetaPlus zeta potential analyzer using Zeta Analysis software (Brookhaven Instruments Corporation, Holtsville, NY, USA). Measurements were performed in 30 mM MES buffer with a pH of 5.99.

Thermogravimetric Analysis (TGA): TGA was performed in a TA Instruments TGA Q-500 thermogravimetric analyzer, using Advantage for Q series v2.5.0.256 and Thermal Advantage v5.4.0 software (New Castle, DE). The temperature was ramped under nitrogen atmosphere at a rate of 10 °C/min from room temperature to 600 °C (700 °C for IONP-OA), with air being introduced at 550 °C.

X-ray photoelectron spectroscopy (XPS): A silicon wafer was washed with acetone, methanol, and isopropyl alcohol. The purified pellet of IONP-DOPAC was re-suspended and added in small amounts to the wafer, which was then dried under nitrogen. XPS measurements were performed on a Thermo Scientific K-Alpha X-ray photoelectron spectrometer, using Thermo Advantage v5.962 software (Waltham, MA). The X-ray

was Al-Kα (1486.7 eV), at a spot size of 400 μm. The plate was washed and the sample was plated in 20 μL aliquots. High resolution Fe spectra were collected at 150 keV pass energy and 50 ms dwell time over 3 scans.

Results

In addition to considerations of time and quantity of material produced, the most important factors in determining the best method for ligand exchange are oxidation state, colloidal stability, particle coverage, and nature of ligand exchange. To inform this process, FT-IR (Fourier-transform infrared spectroscopy) evaluates the nature of coverage, while TGA (thermogravimetric analysis) and zeta potential measurements assess the extent of coverage. Similarly, XPS (x-ray photoelectron spectroscopy) examines the oxidation states.

To begin, comparing the FT-IR spectra for each method allows for an initial distinction to be made between the nature of the coverage of the IONP products. Particularly interesting are the two peaks at 2850 cm^{-1} and 2920 cm^{-1}, whose intensities vary between methods.

The peak associated with Fe-O bonds also varies between methods. The FT-IR spectrum of the twice-sonicated IONPs shifts to 573 cm^{-1}. Meanwhile, similar peaks present in the spectra of the twice-sonicated product and the shaking product both appear at 571 cm^{-1} (Fig. 3).

The aromatic C-H peaks at about 1150 cm^{-1} and 1117 cm^{-1} stem from the DOPAC ligand, qualitatively confirming ligand exchange. While these peaks show significant variation between preparations, their intensities and shapes do not vary. Their location changes most commonly between trials of the twice-sonicated IONPs.

Next, examining the TGA (thermogravimetric analysis) results will shed light on the extent of the IONPs' coverage. To begin with, organic material comprises 19.79% of the starting material, IONP-OAs. Following ligand exchange, the IONPs that undergo shaking have 9.043% organic material, the once-sonicated IONPs have 6.880%, and the twice-sonicated IONPs have 7.597% (Fig. 4).

Next, zeta potential measurements allow the evaluation of each method's colloidal stability. The zeta potential measurements are performed in 30 mM MES buffer with a pH of 5.99 and the results are displayed in Table 1.

Finally, the oxidation state of the iron within the nanoparticles is observed through XPS (x-ray photoelectron spectroscopy) (Fig. 5). The Fe 2p-3/2 peak values for each method are compared with each other and with the original IONP-OA (Fig. 5) (Fe 2p-3/2 refers to the specific state of iron that is relevant for these particular reactions). The Fe 2p-3/2's binding energy for IONP-OA peaks at 711.3 eV, while the once-sonicated product, the twice-sonicated product, and the shaken product peak at 711.6 eV, 711.5 eV, and 711.5 eV, respectively.

Discussion

The FT-IR results are significant in their qualitative confirmation of successful ligand exchange. Each method – sonication once, sonication twice, and shaking – is confirmed by FT-IR to have resulted in successful ligand exchange. With this initial confirmation that OA and DOPAC have indeed been exchanged, the relative success of each method can be determined. However, the variations in peak intensities and fingerprint regions between each method require further investigation to explain. Therefore, FT-IR does not indicate a most successful method, but rather it confirms that ligand exchange has occurred as expected.

The two peaks at 2850 cm^{-1} and 2920 cm^{-1} (Fig. 3) are associated with the C-H stretching vibrations of the OA. These peaks vary in intensity between methods, and are most intense in the one-time-sonicated IONPs' spectrum. More intense peaks in this region are associated with residual OA on the surface of the IONPs. The spectrum of the twice-sonicated IONPs is shifted to 573 cm^{-1}, while the other two are at 571 cm^{-1}. This may indicate oxidation of the iron, as oxidation leads to a more positive charge,

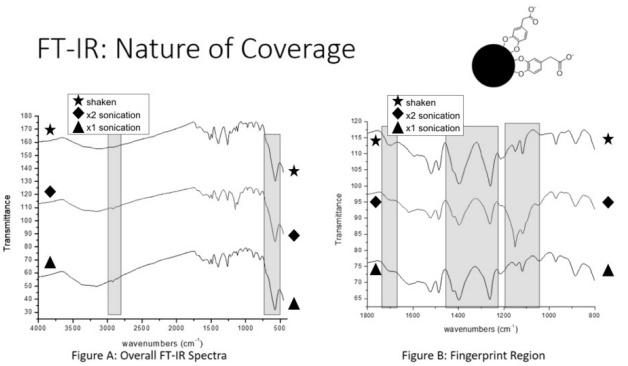


Fig. 3. FT-IR Comparison of Ligand Exchange Methods.

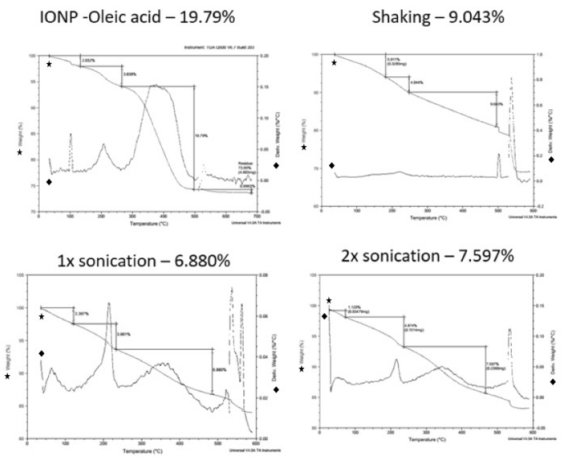


Fig. 4. Results of TGA Analysis Indicating Percentage of Organic Material on Surface.

Technique	Mobility (10^{-8} m^2 / s*V)	Zeta Potential (mV)
Sonication x1 (0.05 mg/mL)	-1.93	-24.75
Sonication x2 (0.05 mg/mL)	-2.49	-31.87
Shaking (0.15 mg/mL)	-2.42	-30.94

Done in 30 mM MES buffer pH 5.99

Table 1. Zeta Potential Measurements.

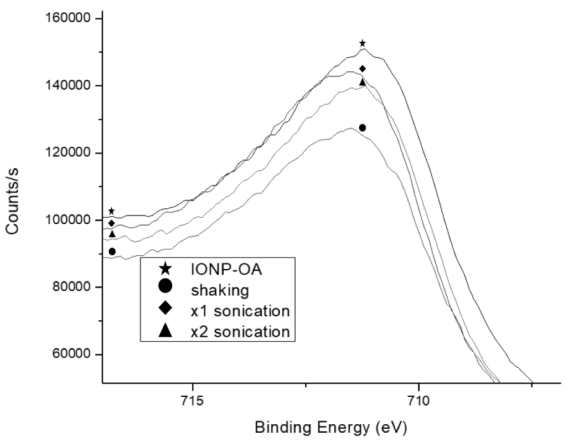


Fig. 5. XPS Comparison of Ligand Exchange Methods.

which results in higher vibrational frequency due to stronger iron-oxygen bonds. This could explain a shift to a higher wavenumber.

While this data confirms that each method results in some ligand exchange, the presence of OA peaks on each spectrum indicates that no method results in complete exchange. The once-sonicated sample has the most intense OA peaks, showing that it results in the least amount of ligand exchange. As such, the once-sonicated method may be the least successful of the three in terms of DOPAC coverage. The possibility of IONP oxidation for the twice-sonicated method, indicated by the shift from 571 cm^{-1} to 572 cm^{-1}, is further investigated through XPS.

Moving on, TGA determines the percentage of DOPAC on the surface of the IONPs, a significant factor in evaluating which method is most useful. The higher the percentage of DOPAC on the surface of the IONPs, the more effective is the method.

Upon initial examination, it appears that the shaking exchange results in the highest percentage of DOPAC on the surface. However, the sudden drop at around 500°C (Fig. 4) is the result of a technical difficulty. The shaking result cannot be relied upon and must therefore be redone.

It can, however, be concluded that the twice-sonicated IONPs are composed of a higher percentage of DOPAC than the once-sonicated IONPs. This indicates the possibility that the double sonication method is superior to the single sonication method.

Now that the extent of coverage has been analyzed through FT-IR and TGA, the nature of coverage is analyzed through zeta potential measurements. These measurements reveal the relative colloidal stability of each method. Colloidal stability is related to the charge of the surface of the nanoparticles, which is indicative of their aqueous stability. Aqueous stability is a highly significant characteristic of IONPs, enabling biomedical and other applications[2].

Here, the products of all the tested methods display zeta potential measurements surpassing the 20 mV threshold, which indicates particle stability (Table 1).(11) To be more specific, the twice-sonicated product has the highest zeta potential, followed by the shaking product, followed by the once-sonicated product. However, as the measurement of the shaking product is performed at a different concentration from the sonication products, a reliable, direct comparison cannot be made. Therefore, between -30.94 mV for the shaking and -31.87 mV for the double sonication, it is unclear which is truly more stable. Comparison between the measurements of the sonicated products, both of which are conducted at the same concentration, indicates that the twice-sonicated product is more aqueous stable. This finding, combined with the results of the TGA, suggests that the twice-sonicated product may be preferable over the once-sonicated product.

Now that coverage and colloidal stability have been compared, the last factor to evaluate is the iron's oxidation state in the IONPs. As the IONPs are meant to be involved in subsequent tests and reactions, it is important to evaluate the oxidation state of the iron, as it will affect the IONPs' behaviour.(12)

According to the results of the XPS, the surfaces of the exchange products of each method are oxidized with respect to the starting IONP-OA. The peak values of the products are all greater than that of the oleic acid, which indicates a higher oxidation state; while the oleic acid peak appears at 711.3 eV, the once-sonicated and shaking products appear at 711.5 eV, and the twice-sonicated product appears at 711.5 eV. It is unclear, however, whether any particular method causes significantly more oxidation of the particles' surfaces. While there is a shift in binding energy before and after the ligand exchange, as evidenced by the difference in binding energy between the oleic acid sample and the rest of the samples, the peaks of the products of each method do not significantly vary from each other at 711.5 eV and 711.6 eV (Fig. 5). Therefore, the XPS results do not indicate whether any exchange technique causes more IONP oxidation than the others. Further trials and peak deconvolution are required to obtain a better understanding of the oxidation state of the iron in the XPS spectra.

Conclusion

Previous studies have investigated the relationships between ligands' chemical structures and the nature of their binding on magnetic IONPs. (9) This study further explores possibilities in ligand-design and ligand-exchange strategies, seeking to determine a standardized method for ligand exchange. The standardization of such a method will allow for reliable production and comparison of custom-built IONPs.

According to the FT-IR measurements, all methods result in ligand exchange. TGA measurements indicate that the twice-sonicated product appears to be the most favourable in terms of DOPAC coverage. According to the XPS measurements, all methods are comparable in terms of oxidation of the iron within the IONPs. It is important to underline that the shaking sample cannot be directly compared to the sonication methods due to inaccuracies of TGA and zeta potential measurements. Additionally, zeta potential measurements should be combined with sizing data, through dynamic light scattering or nanoparticle tracking analysis; this will enable confirmation of colloidal stability.

Further study to determine the best method of ligand exchange would benefit from larger data sets of FT-IR, TGA, zeta potential, and XPS. Zeta potential measurements must be repeated for the shaken product at a concentration that matches the other products, so that a direct comparison may be made. TGA must also be repeated for the shaking product to eliminate possible inaccuracies resulting from technical difficulties encountered in the measurement discussed above. It would additionally be useful in future to repeat the zeta potential measurements in different buffers to explore whether the nanoparticles' colloidal stability varies with the buffer in which they are suspended. Similarly, it would be useful to perform a comparison of these methods as they apply to ligands other than DOPAC, such as dopamine. This will allow for a conclusion that would apply to a wider variety of ligands. Additional testing should also include investigation of the oxidation states of the iron in each method's product. It may be useful to fit the $Fe^{3+/2+}$ octahedral and tetrahedral XPS peaks, in order to examine the ratio. This may provide information about changes in the crystal structure, magnetic properties, and oxidation. Furthermore, use of selected area electron diffraction (SAED), combined with transmission electron microscopy (TEM), may provide useful information about the crystallinity of the various products. It may also provide useful information relating to sizing and morphology of the nanoparticles. Lastly, further study should include comparison of supernatants, as well as FT-IR spectra of twice-sonicated products in methanol and other solvents. This may be useful for exploring the reason for loss of IONPs in the washing step following centrifugation after the second DOPAC addition. It is possible that this loss of product may be minimized or eliminated if a different solvent is used.

Through comparison of various methods of ligand exchange for superparamagnetic aqueous stable IONPs, this study comes one step closer to the drug discoveries and water treatments of the future. Through the standardization of a method of ligand exchange, the enormous potential for biomedical and other advancements can begin to be actualized.

Acknowledgements

Thank you to Dr. Blum and the entire Blum group for their guidance and support, especially to Yifan Ling. Thank you to Shoronia Cross for helping me develop this project and for guiding me every step of the way, from mastering the procedure to data analysis to presentation.

References

1. Zhang JL, Srivastava RS, Misra RDK. Core–Shell Magnetite Nanoparticles Surface Encapsulated with Smart Stimuli-Responsive Polymer: Synthesis, Characterization, and LCST of Viable Drug-Targeting Delivery System. Langmuir. 2007 April 27; 23 (11): 6342-6351.
2. Gupta AK, Gupta M. Synthesis and surface engineering of iron oxide nanoparticles for biomedical applications. Biomaterials. 2004 December 1; 26 (18) 3995-4021.
3. Ling D, Hyeon T. Chemical Design of Biocompatible Iron Oxide Nanoparticles for Medical Applications. Small. 2013 May 1; 9 (9-10): 1450-1466.
4. Laurent S, Forge D, Port M, Roch A, Robic C, Elst L, Muller R. Magnetic Iron Oxide Nanoparticles: Synthesis, Stabilization, Vectorization, Physicochemical Characterizations, and Biological Applications. Chemical Reviews. 2008 June 11; 108 (6): 2064-2110.
5. Lakshmanan R, Okoli C, Boutonnet M, Järås S, Rajarao G., Effect of Magnetic Iron Oxide Nanoparticles in Surface water Treatment: Trace Minerals and Microbes, Bioresource Technology. 2012 December 29; 129: 612-615.
6. Korpany K, Mottillo C, Bachelder J, Dong P, Trudel S, et al. One-Step Ligand Exchange and Switching From Hydrophobic to Water-Stable Hydrophilic Superparamagnetic Iron Oxide Nanoparticles by Mechanochemical Milling. Chemical Communications. 2016 Jan 11; 52: 3054-3057.
7. Yuen A, Hutton G, Masters A, Maschmeyer T. The Interplay of Catechol Ligands with Nanoparticulate Iron Oxides. Dalton Transactions. 12 Jan 2012; 41: 2545-2559.
8. Nagesha D, Plouffe B, Phan M, Lewis L, Sridhar S, Murthy S. 2009 March; 105 (7): 105-107.
9. Korpany K, Majewski D, Chiu C, Cross S, Blum A. Iron Oxide Surface Chemistry: Effect of Chemical Structure on Binding in Benzoic Acid and Catechol Derivatives. Langmuir. 2017 Feb 18; 33 (12): 3000-3013.
10. Korpany K, Habib F, Murugesu M, Blum A. Stable Water-Soluble Iron Oxide Nanoparticles Using Tiron. Mat Chem and Phys. 2012 Oct 7; 138: 29-37.
11. Lu GW, Gao P. Chapter 3 - Emulsions and Microemulsions for Topical and Transdermal Drug Delivery A2 - Kulkarni, Vitthal S. Handbook of Non-Invasive Drug Delivery Systems. Boston: William Andrew Publishing; 2010. p. 59-94.
12. Geng C, Ye S, Neese F. Does a Higher Metal Oxidation State Necessarily Imply Higher Reactivity Toward H-atom Transfer? A Computational Study of C-H Bond Oxidation by High-Valent Iron-Oxo and -Nitrido Complexes. Dalton Transactions. 2014; 43(16): 6079-86.

Submitted: 01/14/18

Research Article

[1]Department of Anatomy & Cell Biology, Mcgill University, Montreal, QC, Canada
[2]Department of Physiology, McGill University, Montreal, QC, Canada
[3]McGill University Health Center Research Institute, Montreal, QC, Canada

Keywords

Planar cell polarity (PCP), Vangl2 gene, polycystic kidney disease (PKD), Apical Constriction (AC), Convergent Extension (CE)

Email Correspondence

ida.derish@mail.mcgill.ca

Ida Derish[1,3], Jeremy Lee[2,3], Sima Babayeva[3], Elena Torban[3]

The Role of Core Planar Cell Polarity Vangl2 Gene in the Renal Tubule Development in Mice

Abstract

Background: Polycystic kidney disease (PKD) is a common kidney disease that affects the development and maintenance of renal tubules, leads to cyst formation, and often progresses to end-stage kidney disease. It has been postulated that defective planar cell polarity (PCP) signaling contributes to initiation of cyst formation in PKD via controlling both convergent extension (CE, a process of directional cell movements) and oriented cell division (OCD, a process of directional cell divisions during tubular elongation post-natally). Indeed, mutations of the key PCP gene, Van Gogh-like 2 (Vangl2), lead to abnormal renal tubules in murine embryonic kidneys, correlating with the original postulate.

Methods: In order to further understand the influence of the Vangl2 gene on renal morphogenesis and cystogenesis, control and Vangl2 mutant embryos—as well as post-natal Vangl2 mice with conditional excision of the Vangl2 gene in renal collecting tubules—were generated, then analyzed using immunostaining and fluorescence microscopy.

Results: Our results show that Vangl2 plays a role in CE and apical constriction (AC) during embryonic stage of tubulogenesis. Compared to control animals, mutant Vangl2$^{\Delta/\Delta}$ and conditional Vangl2$^{\Delta/CD}$ embryos displayed: i) a significant dilation in the diameter of renal tubules seen as an increased tubule cross-section area and a larger number of cells per cross-section; and ii) changes in cell shape indicative of defective AC. Surprisingly, post-natal mice showed virtually no difference in any of these aspects comparing to control mice, suggesting that other pathways may compensate for the lack of PCP signaling in maintenance of the tubule architecture.

Limitations: a) The analysis of the renal tubules at the specific time points does not account for the dynamics of tubular movement and growth in real time; b) a mechanistic and morphological distinction between mice and humans may exist in the renal collecting duct tubules, pertaining to the Vangl2 gene's influence in the PCP pathway; and c) the degree of mosaicism resulting from the gene excision by Cre-recombinase may correlate with the severity of the phenotype.

Conclusion: We conclude that the PCP pathway is required for normal tubule development during embryogenesis. Our results, however, indicate that the cystogenesis seen in PKD postnatally may not be directly attributed to the disrupted PCP signaling, and requires the derangement of additional pathways.

Introduction

Polycystic kidney disease (PKD) is characterized by the formation and progressive enlargement of cysts in the kidney, contributing to an impairment of renal function and eventual end-stage kidney disease.(1) Patients with autosomal dominant polycystic kidney disease (ADPKD), the most common form of PKD, experience a significantly reduced quality of life, and most require extensive therapy, dialysis, and kidney replacement by the age of 55-60.(2) Indeed, ADPKD is the most common potentially lethal single-gene disorder, occurring on average in 1 per 600 PKD cases worldwide.(2-4) There is presently no cure for PKD, which highlights the necessity of research on the mechanisms underlying the development and maintenance of renal tubules.

Formed through mesenchymal-to-epithelial transition and branching morphogenesis, kidney tubules are integral to the proper hemodynamics in the body.(5) Upon its formation, the ureteric bud (UB) receives signals from the mesenchyme, which dictates the branching of UB into structures such as the ureter and, more pertinent to this article, the system of collecting duct tubules of the kidney.(6) In order to promote the branching and tubular growth/elongations, the cells undergo organized, directional intercalation during embryonic development, known as convergent extension (CE), and oriented cell division (OCD), which, in rodents, occurs in the first two weeks after birth.(7,8) Cell proliferation within the tubules is therefore directly linked to the proper elongation of the tubules and an establishment and the maintenance of tubular diameter.

In effect, it is critical for cells to exhibit polarity, as they migrate and differentiate into various tissues, including renal tubules. Throughout several embryonic processes, two polarity axes play a role in defining the proper orientation of epithelial cells. The first, apical-basal polarity, refers to the orientation of the cells perpendicularly to the basal lamina, which demonstrates adhesive properties that signal for the formation of an apical surface at the opposite end of the cell. The second axis of polarity, known as planar cell polarity (PCP), defines cell polarity within a tissue axis (e.g. along a tubule) perpendicular to the apical-basal polarity.(9) Typical examples of PCP are the patterns of *Drosophila* wing hairs and stereocilia in the inner ear of mammals.(10,11) However, it is the role of the PCP pathway in the formation of a uniform kidney tubule diameter that is of greatest concern in this paper.

In 2006, Fischer et al. detected a randomized OCD in two independent postnatal rodent models of polycystic kidney disease. This led others to postulate that defective PCP signaling is causative of cystic initiation, although, until recently, extensive research has not been done to investigate a potential mechanism.(12,13) Mutations of key components in the PCP pathway have been linked to a renal cystic phenotype; for example, an increased tubule diameter in embryonic kidneys of Fat4 (PCP-gene) homozygous mutants was attributed to OCD disruption.(14) Fat4-/- mice

carrying one disrupted allele of Vangl2 exhibited an exacerbated dilatation of renal tubules comparing to Fat4-/-, particularly in the medullary zone that mostly consists of collecting ducts bundles.(14) One recent paper describes the genetic interactions between known PKD-related gene, Pkd1, and Vangl2 in the brain ventricles.(15) Loss of Pkd1 affects planar orientation and position of basal bodies in the multiciliated ependymal cells in the brain; in compound heterozygous mice for Pkd1 and Vangl2, this phenotype was exacerbated. Therefore, there is some evidence that disruption of PCP gene Vangl2 may contribute to PKD, including genetic interaction between the two pathways.

In *Xenopus*, Vangl2 has specifically been linked to the regulation of apical constriction (AC) during gastrulation, an early morphogenetic process which facilitates the proper arrangement of cell layers in the embryo.(16) Ossipova et al. showed that loss of Vangl2 affected cell wedging and recruitment of specific acto-myosin regulators that drive the shortening of the apical surface length (apical constriction).(16) Despite the reports of dilated tubules in embryonic Vangl2 mutant kidneys, the process by which Vangl2 regulates tubular diameter and cystogenesis remains unknown.

In the research presented here, we studied the role of Vangl2 in renal tubulogenesis during embryonic stages, as well as during early postnatal stages. From our results, we report that a loss of Vangl2 leads to tubular dilatation and cystic transformation during embryonic kidney development via the control of the processes of convergent extension and apical constriction. However, architecture of tubules in postnatal mouse kidneys is not affected by the loss of Vangl2. Thus, we conclude that the relationship between defective PCP signaling and PKD is incorrectly labelled as cause-and-effect.

Materials and Methods

Generation of Vangl2 mutant mice

We previously reported the Vangl2 mouse with a Floxed allele which contains LoxP sites flanking Exon 4 (Vangl2$^{FL/FL}$).(17) The generation of the Vangl2 mice with a ubiquitous excision of Exon4 (Vangl2$^{\Delta/\Delta}$ null allele) was previously described.(17) For this study, Vangl2$^{\Delta/\Delta}$ mutant mice were generated by brother–sister mating. These mice are embryonically lethal at around embryonic day (E) 17.5-18.5. Age matched Vangl2$^{FL/FL}$ embryos were used as controls. Hoxb7-Cre mice (expressing Cre-recombinase transgene in the collecting duct tubule (18)) were kindly provided by Dr. Carlton Bates (University of Pittsburg, PA). To generate Vangl2 mice with a conditional excision of Vangl2 in the collecting ducts, series of breedings were undertaken to obtain Hoxb7-CreV2$^{\Delta/CD}$ (conditional mutant Vangl2$^{\Delta/CD}$) mice. Vangl2$^{FL/+}$ mice were used as controls, with intact Floxed alleles due to the lack of Cre allele. The mouse DNA was isolated from a tail biopsy and was analyzed by PCR. All experiments with Vangl2$^{\Delta/\Delta}$ and a matching control were conducted on E17.5 embryos. Post-natal mice at P7 of genotypes Vangl2$^{\Delta/CD}$ and Vangl2$^{FL/+}$ were generated, and the kidneys were excised and preserved for further analysis. All animal work was conducted per the Canadian Animal Care Guidance with approval by the Animal Care Committee, McGill University.

Sectioning of Tissues

Embryonic E17.5 and P7 tissues were fixed overnight in 4% Paraformaldehyde/Phosphate Buffered Saline, pH 7.4, at 4°C. Then, they were dehydrated in various concentrations of ethanol/Phosphate Buffered Saline, embedded in paraffin (McGill University Health Center Research Institute histology service) and sectioned at 4 μm on the microtome. Each section was viewed on a light microscope to assess the depth of sectioning. The sections were mounted on slides (2 per slide) and dried in the incubator overnight at 37°C.

Immunofluorescence Staining

The E17.5 and P7 sections were de-paraffinized by heat at 55°C in a dry incubator and treated with xylene, according to standard protocol. The sections were rehydrated in the ethanol solutions of progressively decreasing concentrations in the Phosphate Buffered Saline, pH 7.4. The epitope retrieval procedure was performed by immersing slides in the 10 nM Na-Citrate buffer which was pre-heated at 100°C (brought to a boil) (Vector Laboratories, CA). The slides were then incubated in a microwave for 20 min at maximum heating. Special care was taken to ensure that the boiling buffer was covering the sections at all times. The immunostaining was conducted as follows: first, the tissues were blocked overnight at 4°C with a solution containing 3% Bovine Serum Albumin, 5% Normal Goat Serum, 5% Normal Donkey Serum, 0.1% TritonX, and 0.05% Sodium Dodecyl Sulphate in the PBS (pH 7.4). The tissues were then stained with anti-Calbindin (the specific marker of collecting ducts, 1:500, D-28K) and anti-E-Cadherin (a marker of epithelial cells, 1:500) antibodies followed by the incubation with secondary donkey anti-mouse IgG Alexa 546 and goat- anti-rabbit IgG Alexa 488 antibodies. The nuclei were visualized by staining with 4',6-diamidino-2-phenylindole (DAPI). Fluorescence microscopy was performed using the Zeiss AxioObserver Z1 inverted fluorescence microscope.

Morphological analysis

Low Magnification Analysis: Low magnification images were compared by looking at three renal morphological features: a) the collecting duct tubules; b) the cortico-medullary definition; and c) the overall size of the kidney, in relation to the scale bar. Dilated tubules were defined by their larger lumen, whereas cystic tubules displayed an aggravated dilatation and were lined by a flattened and disorganized epithelium.

Convergent Extension Analysis: CE analysis was performed on the images using the Zeiss ZEN disk 2012 (Blue Edition) program (Zeiss, Germany). Analysis consisted of measuring the cross-sectional area of the tubule, as well as its width. We defined the 'width' as the longest axis seen in the tubular cross section. The perpendicular axis to this width was also measured and the ratio of the two diameters was calculated. Only the cross-sections with measurements within 10% of each other were qualified as "perfectly" transverse and used for the CE analysis. For E17.5 embryos and P7 mice, a minimum of 20 perfect transverse structures were used for every animal to calculate the tubular area and width. The cross-sections used for the CE analysis were also used to count the number of nuclei. Four embryos and three P7 mice per genotype were analyzed.

Apical Constriction Analysis: Apical constriction (AC) analysis was performed on the transverse "perfect" tubule sections of the E17.5 kidneys by measuring the apical and the basal lengths of each cell (visualized with anti-Cadherin antibody). These apical and basal lengths were defined as the size of the most apical and basal surfaces of the cell, spanning from one lateral membrane to the other. The ratio of apical-to-basal lengths was calculated for each cell. The measurements were conducted on at least 60 transverse tubule sections in 4 embryos per genotype.

Statistical Analysis

To test for significant differences (p-value) between the results, the Student's t-test with an equal two-tailed distribution and an unequal variance was used. A p-value smaller than 0.05 was considered as statistically significant. Standard mean errors are shown in each bar graph.

Results

Analysis of Embryonic (E17.5) Mouse Kidney

In order to ascertain the role of Vangl2 in embryonic renal tubular development, embryonic day (E) 17.5 sections of the control (Vangl2$^{FL/FL}$), null mutant (Vangl2$^{\Delta/\Delta}$, with ubiquitous loss of Vangl2 in all cells), and conditional (Vangl2$^{\Delta/CD}$ Cre+) embryos were analyzed. Embryonic Vangl2$^{\Delta/\Delta}$ kidneys displayed the following abnormalities: i) hypoplasia (at low magnification, mutant maximal kidney sections fit into the 'image frame', whereas a larger number of images were required to capture the entire area of the maximal control kidney, indicative of hypoplasic mutant kidneys); ii) a disorganized tubular architecture with a profound loss of the medullary zone; and iii) the presence of dilated and cystic

tubules (Fig. 1). Vangl2$^{\Delta/CD}$ kidneys also displayed an abnormal renal morphology, albeit the degree of abnormalities was more moderate. Control kidneys had a clearly demarcated cortico-medullary division, and did not display any dilated tubules or cystic structures.

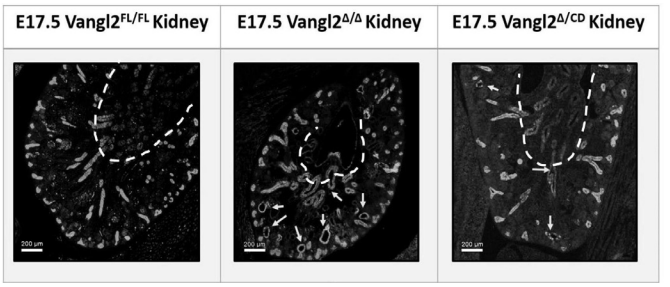


Fig. 1. Morphology of E17.5 control Vangl2FL/FL, null Vangl2$^{\Delta/\Delta}$, and conditional Vangl2$^{\Delta/CD}$ (HoxB7CreV2$^{\Delta/CD}$) kidneys (50X magnification). Kidneys immunostained with anti-Calbindin (green, collecting duct marker), anti-E-Cadherin (red, epithelial tubular cells marker) antibodies and DAPI (blue, nuclei marker) stained. Dotted lines indicate the cortico-medullary definition. Vangl2$^{\Delta/\Delta}$ kidneys are hypodysplastic, with dilated cystic collecting ducts (arrows). Vangl2$^{\Delta/CD}$ kidneys feature dilated tubules and occasional cystic phenotype.

During embryogenesis, renal tubule formation has been shown to be controlled by convergent extension, regulated by PCP signaling.(19) To assess CE (regulated by the PCP pathway), we measured the cross-sectional area and additionally counted the average number of cells in the "perfect" collecting duct cross sections (differences of two perpendicular axes are less than 10%). We found that E17.5 Vangl2$^{\Delta/\Delta}$ and Vangl2$^{\Delta/CD}$ renal tubules had a cross-sectional area that was significantly larger than that of the controls, indicative of the tubular dilatation (Fig. 2). We likewise observed that the average number of cells surrounding the tubular lumen (visualized by the DAPI-stained nuclei) was statistically higher in both null and conditional kidneys, which was another manifestation of the failed CE that led to an excessively dilated tubular phenotype (Fig. 2).

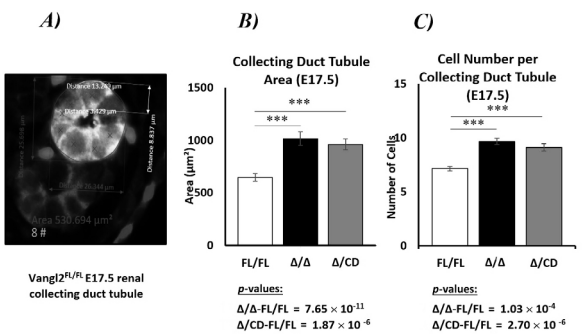


Fig. 2. Convergent extension (CE) analysis in control, null, and conditional renal collecting ducts in embryonic E17.5 kidneys. A) Perfect cross-section of a collecting duct tubule at 40X (perpendicular measurements in red); B) Statistical analysis of the collecting ducts area.; C) Statistical analysis of the cell number per collecting duct (Vangl2$^{FL/FL}$; N=87. Vangl2$^{\Delta/\Delta}$; N=85. Vangl2$^{\Delta/CD}$; N=113). ***p<0.001.

To analyze apical constriction, the same tubular cross sections chosen for convergent extension measurements were studied. The apical constriction describes the process of cellular 'wedging', where the apical surface of the cell is constricted due to the recruitment of specific proteins to that surface. This allows for a tighter cell packaging and the reduction of the diameter in developing tubules. The collecting ducts (identified by Calbindin staining) were stained with anti-E-cadherin to delineate lateral borders of each cell; the apical and basal lengths were measured and the apical-to-basal ratio was calculated. We observed that this ratio was larger in both null and conditional mutants (Fig. 3), indicating that the cells were more cuboidal. Although significantly different from controls, we consistently found that the tubular phenotype in conditional mice was milder than in null mutants, including fewer cells in the tubular cross-sections, as well as less profound changes in cell shape.

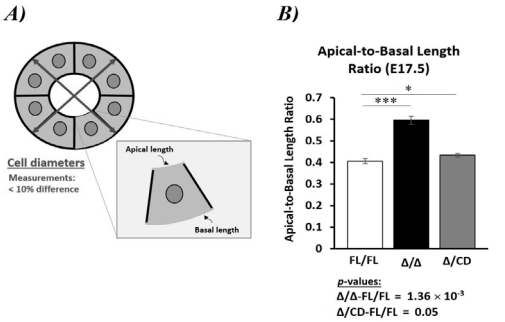


Fig. 3. Apical constriction (AC) analysis in control, null and conditional renal collecting ducts in embryonic E17.5 kidneys. A) Depiction of AC measurements: apical and basal surface lengths were measured in "perfect" tubule cross-sections and the ratios were calculated for each cell. B) Statistical analysis of AC in collecting ducts (Vangl2$^{FL/FL}$; N=87, Vangl2$^{\Delta/\Delta}$; N=85, Vangl2$^{\Delta/CD}$; N=113). ***p<0.001.

In summary, our results indicate that during the embryonic stage, defective CE and AC lead to dilated collecting ducts as well as cyst formation in the Vangl2 deficient tissues.

Analysis of Post-natal (P7) Mouse Kidneys

Vangl2$^{\Delta/\Delta}$ mice are embryonically lethal. To circumvent the embryonic lethality and to ascertain whether the PCP deficiency leads to cystogenesis post-natally, Vangl2$^{\Delta/CD}$ mutant mice were generated through Cre-recombinase knockout of the Vangl2 gene in collecting ducts. Cre-negative Vangl2$^{FL/+}$ mice were used as a control.

Morphological analysis of post-natal (P7) Vangl2$^{\Delta/CD}$ and control Vangl2$^{FL/+}$ kidneys revealed that, surprisingly, mutant kidneys were similar in size to control kidneys and lacked any discernable cystic phenotype (Fig. 4). Furthermore, when the average cross-sectional area of collecting ducts was analyzed at 400X, we observed no significant difference in the diameter width between the mutant and control kidneys, indicating a lack tubular dilatation.

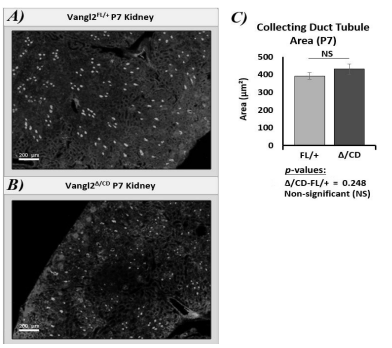


Fig. 4. Morphology of the postnatal day 7 (P7) Vangl2Δ/CD and Vangl2 FL/+ kidneys. A,B) Kidneys were immunostained with anti-Calbindin (green) and anti-ECAD (red), and stained with DAPI (blue), and imaged at 50X magnification; C) Statistical analysis of the cross-sectional area in the P7 Vangl2$^{\Delta/CD}$ (N=87) and Vangl2 $^{FL/+}$ (N=69) renal mouse collecting duct tubules.

Discussion

In this study, we have shown that the PCP gene Vangl2 regulates the processes of convergent extension and apical constriction in the renal tubules during embryonic kidney development. These observations are novel and suggest a functional conservation of the PCP gene regulation of CE and AC in various tissues. However, we found that the post-natal P7 mutant kidney phenotype is indistinguishable from the control kidneys with regard to the size and shape of the collecting duct tubules. This observation is

rather unexpected. For over 10 years, the research community has strongly believed that the PCP pathway controls tubular diameter and, when deficient, contributes to the tubular dilation and cyst appearance featured in polycystic kidney disease.(12,13) Our observations are consistent with the notion that pathways others than PCP (e.g. the ones that are regulated by the ADPKD genes Pkd1 and Pkd2) are likely responsible for the control of tubular diameters during post-natal phase of tubular elongation and the onset of PKD.(20)

Kunimoto et al. recently reported a similar lack of cystic transformation postnatally in the double Vangl2/Vangl1 and Fz3/Fz6 double homozygous mutants.(8) Through a series of elegant experiments, the authors convincingly showed that the collecting duct tubules in the mutant 16 week-old mice were similar to that of controls, despite some statistically enlarged tubule diameters seen in mutant tissues during the embryonic stage. The authors, however, did not address the mechanisms of tubular dilatation in embryonic tissues nor did they describe cystic structures in the embryonic kidneys.

Both our own observations and those reported by Kunimoto suggest that there might be a switch in the regulating pathways that control the tubular diameter during embryonic kidney development and after birth. We have analyzed E17.5 (a highly statistically significant difference) and P7 (no difference) tissues. We, therefore, surmise that the "switch" occurs somewhere in between. In the future, it would be important to analyze kidneys at the intermediate stages to pinpoint the exact timing of the mechanistic changes that regulate collecting duct tubule diameter.

Our data is corroborated by the study of the inner ear development in the Vangl2 mutant.(21) Copley et al. detected a profound defect in the planar polarity of the stereociliary bundles on the sensory hair cells in the cochlea's organ of Corti in E18.5 mutants. However, 10 days after birth, Copley et al. observed a rescue of the PCP defect and the realignment of the stereociliary bundles. The authors concluded that the refinement process was Vangl2-independent. Our study of renal collecting duct tubules is also indicative of a possibility of such mechanistic switch or a 'refinement process'.

Cells rely on internal signals to intercalate in a highly organized and coherent manner, which Vangl2 mediates through CE and AC. In Vangl2$^{\Delta/\Delta}$ and Vangl2$^{\Delta/CD}$ embryos, this movement was clearly disrupted, as shown by the higher number of cells forming the circumference of the tubule, a larger diameter, and a cystic phenotype. The findings of defective apical constriction and the consequent tubule diameter dilation point towards Vangl2's essential role in renal physiology. The ideal 'wedged cell' shape normally arises from the constriction of the apical side due to the recruitment of proteins which regulate actomyosin and drive constriction of the apical cell surface. This leads to a smaller tubular diameter. The fact that the Vangl2$^{\Delta/\Delta}$ collecting duct cells have a larger apical-to-basal length ratio and a more cuboidal shape means that AC of the renal cells is partially controlled by the PCP pathway. Conditional mice showed a smaller apical-to-basal length ratio than the embryonic mutants, but still displayed a significant difference when compared to Vangl2$^{FL/FL}$. This statistical difference is biologically relevant despite the milder phenotype, since it might still contribute, at least partly, to tubular dilatation consistently seen in conditional embryos. Dysregulation of renal collecting duct tubule diameter might contribute to cystogenesis and as such, a small statistical difference is significant in the context of renal dysfunction and disease.

Our study has some limitations that must be considered. One limitation is that tubular kidney development was not assessed in a dynamic way. Indeed, the analysis involved taking pictures to document the progress (or lack thereof) of cyst formation at the specific time points. However, the process of tubulogenesis features continuous complex movements and 3-dimensional elongation, and our study did not account for this. Moreover, despite the fact that PCP-gene expression is well-conserved between vertebrates, potential mechanistic differences between mice and humans have not been taken into account. Milder phenotype of renal tubules in the E17.5 conditional Vangl2$^{\Delta/CD}$ tissues versus the null Vangl2$^{\Delta/\Delta}$ mutants can be explained by the Cre-recombinase mosaic excision.

Our mouse models paint a convincing image of the Vangl2 effects on renal tubulogenesis. PCP signaling is required for normal tubule development, but defective PCP is not solely culpable for the defective renal tubulogenesis and cytogenesis seen in PKD. Further research on the potential pathogenic signaling pathways leading to tubule dilation and cyst formation postnatally is necessary to fully understand the underlying mechanisms of PKD.

Acknowledgements

We would like to thank our reviewers for their constructive comments and help.

References

1. Harris PC, Torres VE. Polycystic Kidney Disease. Annu Rev Med. 2009; 60:321-337.
2. Harris PC, Torres VE. Polycystic Kidney Disease, Autosomal Dominant. GeneReviews [Internet]. 2002 Jan 10 [Updated 2015 Jun 11]; Available from: https://www.ncbi.nlm.nih.gov/books/NBK1246/
3. Gabow PA. Autosomal dominant polycystic kidney disease. N Engl J Med. 1993 Jul 29; 329(5):332-42.
4. Patel V., Chowdhury R., Igarashi P. Advances in the pathogenesis and treatment of polycystic kidney disease. Curr. Opin. Nephrol. Hypertens. 2009; 18:99–106.
5. Campbell K, Casanova J, Skaer H. Mesenchymal-to-epithelial transition of intercalating cells in Drosophila renal tubules depends on polarity cues from epithelial neighbours. Mech. Dev. 2010;127(7-8):345–357.
6. Constantini, F. Genetic controls and cellular behaviours in branching morphogenesis of the renal collecting system. WIREs Dev. Biol. 2012;1(5):693–713.
7. Lienkamp, SS, Lui K, Karner CM, Carroll TJ, Ronneberger O, et al. Vertebrate kidney tubules elongate using a planar cell polarity-dependent, rosette-based mechanism of convergent extension. Nat. Gen. 2012;44(12):1382–1387.
8. Kunimoto K, Bayly RD, Vladar EK, Vonderfecht T, Gallagher AR., et al. Disruption of Core Planar Cell Polarity Signaling Regulates Renal Tubule Morphogenesis but Is Not Cystogenic. Curr Biol. 2017 Oct 23;27(20):3120-3131.
9. Carroll TJ, Das A. Planar cell polarity in kidney development and disease. Organogenesis. 2011;7(3): 180–190.
10. Maung SMTW, Jenny A. Planar cell polarity in Drosophila. Organogenesis.2011;7(3):165-179.
11. Butler MT, Wallingford JB. Planar cell polarity in development and disease. Nat. Rev. Mol. Cel. Biol. 2017 Mar 15; 18: 375-388.
12. 12. Fischer E, Legue E, Doyen A., Nato F, Nicolas JF, et al. 2006. Defective planar cell polarity in polycystic kidney disease. Nat. Gen. 2006; 38(1):21–23.
13. Happé H, de Heer E, Peters DJ. Polycystic kidney disease: the complexity of planar cell polarity and signaling during tissue regeneration and cyst formation. Biochim Biophys Acta. 2011 Oct;1812(10):1249-1255.
14. Saburi S, Hester I, Fischer E, Pontoglio M, Eremina V, et al. Loss of Fat4 disrupts PCP signaling and oriented cell division and leads to cystic kidney disease. Nat Gen. 2008 Aug;40(8):1010-5.
15. Ohata, S., Herranz-Perez, V., Nakatani, J., Boletta, A., Garcia-Verdugo, J.M., and Alvarez-Buylla, A. Mechanosensory Genes Pkd1 and Pkd2 Contribute to the Planar Polarization of Brain Ventricular Epithelium. J Neurosci. 2015; 35:11153-11168.
16. Ossipova O, Chuykin I, Chu CW, Sokol SY. Vangl2 cooperates with Rab11 and Myosin V to regulate apical constriction during vertebrate gastrulation. Development, 2015; 142:99-107.
17. Rocque BL, Babayeva S, Li J, Leung V, Nezvitsky L, et al. Deficiency of the planar cell polarity protein Vangl2 in podocytes affects glomerular morphogenesis and increases susceptibility to injury. J Am Soc Nephrol. 2015 Mar;26(3):576-86.
18. Zhao H, Kegg H, Gradya S, Truong HT, Robinson ML, et al. Role of fibroblast growth factor receptors 1 and 2 in the ureteric bud. Dev. Biol. 2004 Dec 15; 276(2): 403–415.
19. Karner CM, Chirumamilla R, Aoki S, Igarashi P, Wallingford JB, et al. Wnt9b signaling regulates planar cell polarity and kidney tubule morphogenesis. Nat Gen. 2009 Jul;41(7):793-9.
20. Luyten A, Su X, Gondela S, Chen Y, Rompani S, Takakura A, Zhou J. Aberrant regulation of planar cell polarity in polycystic kidney disease. J Am Soc Nephrol. 2010 Sep; 21(9):1521-32.
21. Copley CO, Duncan JS, Liu C, Cheng H, Deans MR. Postnatal refinement of auditory hair cell planar polarity deficits occurs in the absence of Vangl2. J Neurosci. 2013 Aug 28;33(35):14001-16.

Submitted: 01/13/18

Research Article

[1]Atmospheric and Oceanic Sciences, McGill University, Montreal, QC, Canada
[2]NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, USA

Keywords

Southern ocean, sea ice, mesoscale eddies, polynya, anthropogenic forcing

Email Correspondence

jade.sauve@videotron.ca

Jade Sauvé[1], Carolina Dufour[1], Stephen M. Griffies[2], Michael Winton[2]

Antarctic Sea Ice Trends: Insights from a Suite of Climate Models

Abstract

Background: Antarctic sea ice concentration has been observed to increase from 1978 to 2015, in contrast with the decrease that most climate models show. Here, we aim to examine the respective roles of natural variability and anthropogenic forcing in shaping Antarctic sea ice trend.

Method: To do so, we use the GFDL-CM2 coupled climate model with varying horizontal resolutions in the ocean (1°, 0.25° and 0.10°) that displays a range of behaviours in natural variability with the representation of Weddell Sea polynyas, and different intensities in the decrease of sea ice under climate change.

Results: In the 0.10° model, a sea ice trend of similar sign and magnitude to that observed over the satellite record is found between two occurrences of the Weddell Sea polynya. In the 1° and 0.25° models, which do not simulate any polynya, no equivalent trend of what the satellite record shows is found. Under increasing CO_2 forcing, all models show a surface cooling on a short time scale (years) south of 50°S, followed by a warming on a longer time scale (decades), consistent with the delayed warming mechanism of Ferreira *et al*. (2015). Of all models, the higher resolution model shows the strongest surface warming and decrease in sea ice, suggesting an important role for mesoscale eddies in the response of Antarctic sea ice to climate change.

Conclusion: We conclude that the Weddell Sea polynya is key to the representation of the sea ice trend and that the disagreement between models and observations might partly arise from a desynchronization of the polynya cycles or a too weak natural variability of sea ice in models compared to observations.

Introduction

Despite global warming, Antarctic sea ice extent has been steadily increasing over recent decades and sea surface temperatures (SST) have been decreasing in contrast with the decrease in sea ice cover and the increase of SST in the Arctic.(1) The satellite passive-microwave data from 1978 to 2010 shows an overall increasing trend of 17 100 ± 2300 km^2 yr^{-1} south of the Antarctic Circumpolar Current.(2) The trend in sea ice concentration is not homogeneous around Antarctica however, with some regions even experiencing a strong decrease (e.g. Amundsen-Bellinghausen Sea).(2) In contrast, most models differ from observations by representing a mostly homogeneous decrease in sea ice extent.(3)

The current hypotheses that attempt to explain the observed increase in sea ice can be broadly separated into two categories: the ones caused by natural variability and the ones triggered by forced variability also referred to as anthropogenic forcing.

A study by Polvani and Smith (4) offers evidence that natural variability in sea ice overwhelms the response to anthropogenic forcing by showing that the current trend in Antarctic sea ice is well within the range of simulated trends from preindustrial simulations of Coupled Model Intercomparison Project Phase 5 (CMIP5) models, and that trends induced by anthropogenic forcing are comparatively small. One of the prominent features of sea ice natural variability in models is the appearance and disappearance of open-ocean polynyas that mostly form in the Weddell Sea.(5,6) These ice-free expanses in the otherwise sea ice covered region are caused by an upwelling of relatively warm water from depth and maintained by convection.(7) The first satellite observations of Antarctic sea ice have allowed us to observe an open-ocean polynya in the Weddell sea that lasted from 1974 to 1976 which shows that such events do take place outside from model simulations. (8) During the formation of a polynya, there is a decrease in the Antarctic sea ice extent. As the polynya closes, the Antarctic sea ice extent area increases back to a non-polynya state. Deep and abyssal warming have been reported since the late 1970s and can be partly attributed to a rebound from the 1974-1976 Weddell Sea polynya, which is often referred to as a recovery period.(9) That recovery period might be ending as the record low Antarctic sea ice extent of the Austral winter 2016 as well as the small polynya that opened in the Weddell Sea during 2017 early signs of a polynya.(10)

Anthropogenic forcing has also been posited to be the cause of the current sea ice increase in Antarctica. One theory postulates that increased basal melt of Antarctic ice shelves leads to the production of a cool and fresh surface layer that prevents warmer water from melting sea ice and favours the formation of more sea ice.(11) However, Swart and Fyfe (12) showed that freshening of Antarctic surface waters produces only a small effect on sea ice over the historical period and that a freshening of surface waters fails to reproduce the patterns of sea ice trends in the Southern Ocean. Another possible explanation is that increased downward heat flux from the surface to the deep ocean and increased precipitation minus evaporation leads to increased stratification in the upper ocean and inhibition of the upward flux of heat from warmer water at depth.(1) Alternatively, Holland and Kwok (13) attribute the current sea ice trend to wind changes in the Southern Ocean by demonstrating that local ice-motion is directly related to the local wind trend. They also note that regions with a meaningful increase in northward ice flow present an increase in sea

ice concentration and vice versa.

On the other hand, anthropogenic forcing related to ozone depletion and increase in greenhouse gases has been posited to cause a decrease in sea ice.(14) Indeed, most models forced under an historical scenario simulate a decrease in sea ice.(12) In an attempt to reconcile results from models and observations, Ferreira et al.(15) proposed the delayed warming mechanism to explain the current decrease in SST and increase in sea ice solely due to anthropogenic forcing. This mechanism links the increase in westerly winds, caused by ozone depletion in the stratosphere over Antarctica (16), to the delayed warming of the ocean surface observed in the Southern Ocean. The mechanism involves a two-step response which is illustrated in Fig.1. First, in response to the wind intensification, there is an immediate increase in Ekman advection that produces an initial cooling around Antarctica (Fig.1a.). This cooling then leads to the production of more sea ice, which would explain the current sea ice trend. Then, the slow but persisting response is a warming at all latitude south of 30°S, causing a decrease in sea ice (Fig.1b.). This decrease in sea ice is due to the increased Ekman currents being divergent and causing anomalous upwelling of relatively warm water south of 50°S. (15) Armour et al.(17) suggested that Southern Ocean delayed warming is directly dependent on the timescale of North Atlantic deep waters warming due to the global meridional overturning circulation.

Mesoscale eddies have been found to be key players in the second step of the delayed warming mechanism.(18) These transient ocean features, 10 to 100 km large, are often referred to as the 'weather of the ocean', playing the role of cyclones and anticyclones in the atmosphere. They contain as much as half of the kinetic energy of the ocean and are responsible for transporting and mixing tracers in the Southern Ocean.(19) In particular, eddies are responsible for transporting heat poleward across the Antarctic Circumpolar Current. It has been shown that an increase in westerly winds is followed by an increase in eddy kinetic energy (EKE) with a 2 to 3 years lag.(20) This lag has been attributed to the time it takes surface changes to reach the circulation of the deep ocean. Indeed, the excess energy originating from the increased winds is first stored as potential energy until eddies gradually transfer momentum from the surface to the deep ocean. This increase in EKE is stronger as the resolution of the model is higher. Models which do not resolve eddies will not represent the full temperature response but will be able to show the short-term response of the delayed mechanism.(18)

The purpose of this study is to assess the respective roles of natural variability and anthropogenically forced ocean warming in shaping the Antarctic sea ice trend. First, we consider the role of natural variability and we hypothesize that the recovery from the 1970's polynya can largely explain the current observed positive sea ice trend. Second, we hypothesize that the increase in the westerly winds due to anthropogenic forcing induces a decrease in sea ice on the long term that is augmented by the presence of mesoscale eddies through southward heat transport. As such, we propose that a better representation of the Weddell Sea polynya and of mesoscale eddies in models will allow a more accurate prediction of the sea ice trend. To test these hypotheses, we use a suite of three coupled climate models that differ by the resolution of the ocean component thus allowing us to explore the role of mesoscale eddies in the transport of heat towards the seasonally sea ice covered region. These models also offer a range of behaviours in natural variability through the representation of the Weddell Sea polynya,

Methods

Observational Dataset

We study two variables: sea ice concentration and SST. Sea ice concentration corresponds to the fraction of each observed or modelled grid cell covered by sea ice. It is expressed between 0 and 1. The sea ice extent area, calculated from the former, is the total area covered by sea ice, in km^2, with a sea ice concentration threshold of 0.15 chosen to select which grid cells are included in the overall area of sea ice.

We study the sea ice concentration dataset from the National Snow and Ice Data Center (NSIDC V2) at the National Oceanic and Atmospheric Administration (NOAA). Its source is passive microwave data from satellites. This dataset is computed using two algorithms from the NASA Goddard Space Flight Center (GSFC). The final product uses the highest value from either algorithm for each grid cell.(21) We use monthly values from 1987 to 2015. The limitations of this dataset include a tendency to under-estimate sea ice concentration especially in Antarctic winter. The SST data originates from the Hadley Centre Global Sea Ice and Sea Surface Temperature (HadISST). It consists of monthly values from in-situ observations and adjusted satellite data. The data set extends from 1870 to 2017. Grid cells containing more than 90% of sea ice were set at the freezing point temperature (-1.8 °C). The greatest strength of this dataset is its overall global spatial completeness, though it is less so in the polar regions, especially the Southern Ocean.(22)

The CM2-O Climate Model Suite

We use a suite of three coupled climate models from the Geophysical Fluid Dynamics Laboratory.(23) Ocean, land, atmosphere, sea ice and their interactions are modelled. The models differ only by the horizontal resolution in the ocean. The lowest resolution model is CM2-1deg (1° horizontal resolution) and is the only model of the suite which is run with a mesoscale eddy transport parameterization.(23) CM2.5 has a resolution of 0.25° and CM2.6 has a resolution of 0.1°. A moderate and rich mesoscale eddy fields are resolved in each model respectively. The vertical ocean resolution is 50 levels with a thickness of 10 m at the surface increasing with depth to 210 m. CM2.6 resembles the most observational estimates of dynamic sea level, as expected from its refined resolution (see **Fig. 1** of Griffies et al.(23)).

Each version of the models has two different experiments available for both of which we analyse a period of 80 years. The control experiment is used to investigate the natural variability of the model. The CO_2 concentration is kept constant at a preindustrial level (286 ppm). The perturbation experiment is an idealized climate change scenario where the atmospheric CO_2 concentration undergoes a 1% increase per year. It comprises both the natural variability and the forced response of the system. The difference between the perturbation and control experiments allow us to estimate the response to anthropogenic forcing. Neither the perturbation nor the control experiments correspond to specific years in the historical record. To compare our model output with the observational dataset, we will use years 22 to 50 (352 ppm to 466 ppm) from the perturbation experiment as the atmospheric CO_2 concentration for model year 22 matches the CO_2 concentration in April 1989, which is near the start of the observational dataset.

Results

Evaluation of Models Against Observations

I. Mean State

We start by assessing the realism of the models in representing sea ice extent. Fig. 2 illustrates the seasonal variability of sea ice extent area around Antarctica for observations and the perturbation experiment of the models for the 30 years most similar to observations for CO_2 concentration (for years 1987-2015 of observations and years 22-50 of the models). Models and observations present the same general pattern. Both models and observations reach their lowest value in February and their highest value in September or October. The models present

larger amplitude than observations in their seasonal cycle and show higher sea ice extent area in winter and also slightly lower sea ice extent area in summer. Indeed, we observe a maximum of 21% increase in amplitude between observations and CM2.5 in winter and 84% decrease in amplitude between observations and CM2.6 in summer. However, CM2.6 comes very close to observations for the winter period with only 2.5% larger sea ice extent area than observations. In general, the seasonal cycles are similar to the observational dataset.

We also compare maps of the SST averaged over the whole period of study in order to evaluate the spatial pattern of models against observations (Fig.s not shown here). The general pattern of models and observations is similar with temperature increasing from Antarctica to the Equator. The range of temperature is also the same from -1.8 °C to around 20 °C. Models show more spatial variability which is most likely due to higher resolution for CM2.5 and CM2.6. Overall, the models offer a reasonable comparison to observations.

II. Long-term Trend

Next, we ascertain the differences between the sea ice trends in observations and models. Plotted on Fig. 3 are time series of annual sea ice extent area in Antarctica for observations and the CO_2 perturbation experiment of the models. A linear regression over the complete time period is calculated for each data set with the coefficient of determination R^2 and the linear regression coefficient a. We observe that the models simulate a negative trend while observations show a positive trend. This is not an unexpected result as it is a typical discrepancy between models and observations for sea ice in Antarctica.(3) In addition, the amplitude of the sea ice trend for satellite observations is smaller than for all models (+0.3 vs -0.4, -0.7, -0.8 million km^2 per year for CM2-1deg, CM2.5 and CM2.6 respectively). CM2.5 and CM2.6 show trends of similar magnitude, differing from the sea ice extent trend in CM2-1deg, which does not decrease as much. It is also interesting to note that the decreasing trend for CM2.6 stalls between model years 20 to 50 because of the formation of large polynyas in the Weddell Sea.

Role of the Wedell Sea Polynya on Sea Ice Trend

We now consider the effect of natural variability on the sea ice trend in Antarctica to evaluate our first hypothesis. Fig. 4 shows a time series of the annual average of the sea ice extent area around Antarctica for observations and the control experiment of the models. The control experiment admits no anthropogenic changes in atmospheric CO_2 and ozone concentrations. CM2.6's variability largely differs from that of the two other models due to the simulation of Weddell Sea polynyas. CM2.6 is the only model of the suite that simulates open-ocean polynyas (Dufour et al.(24)). Polynyas form spontaneously in CM2.6 as is the case in many models. These polynyas induce a strong variability in the CM2.6 time series that is more obvious in the control simulation (Fig. 4) than in the perturbation simulation (Fig. 3), although both the control and perturbation experiments admit polynyas. In the perturbation experiment, the variability is due to the superposition of the climate change trend on the natural variability. In Fig. 4, we observe two polynyas, one from years 2 to 30 and one from years 62 to the end of the time series. These polynyas both form in the Weddell Sea, west of the Greenwich meridian like the one observed in the 1970s. The polynyas in CM2.6 are bigger in size than the one observed (2-3x10^5 km^2 for the 1970s polynya and 11x10^5 km^2 for the modelled polynya).(24, 25) At their widest, model polynyas are not completely enclosed by sea ice, like observed in the 1970s, but appear like embayments. Despite these differences, CM2.6 shows a positive trend in sea ice extent area similar in

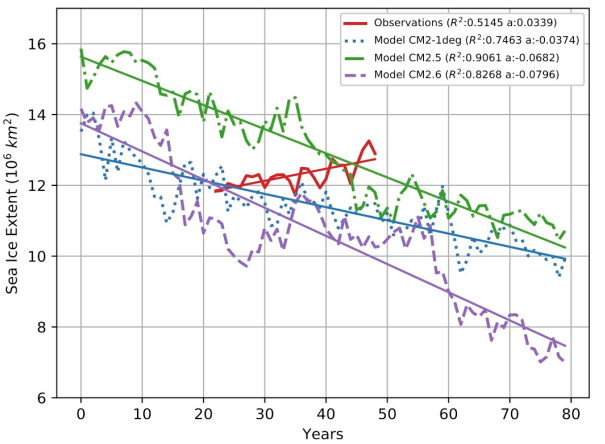


Fig. 3. Time series of annual sea ice extent area around Antarctica for observations for the 1987-2015 period and the perturbation experiment of the models. We calculate the coefficient of determination R2 and the trend a (millions of square kilometers). The time series of observations has been shifted so that the concentration of atmospheric CO2 in the model corresponds to the concentration of the first year of observations

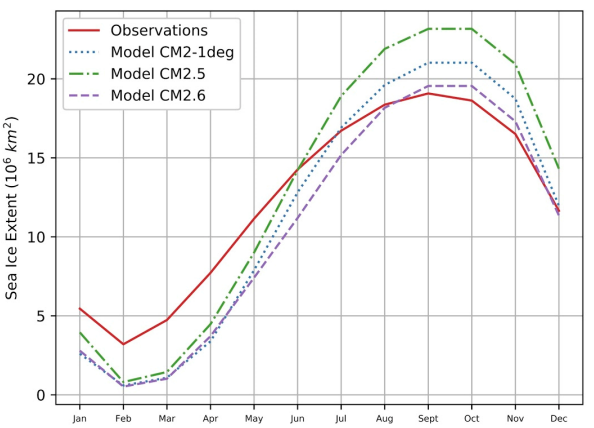


Fig. 2. Seasonal variability of sea ice extend area around Antarctica for observations for the 1987-2015 period and the perturbation experiment of the models.

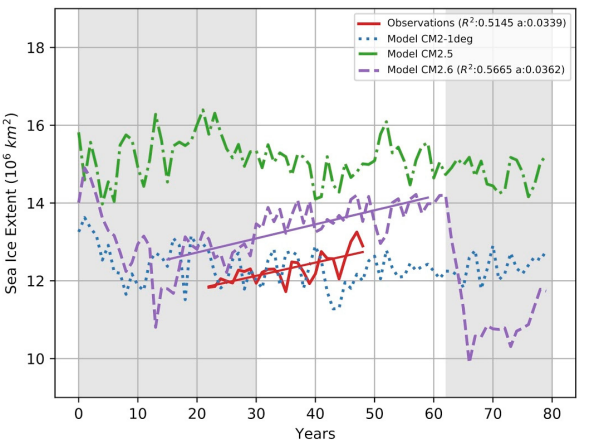


Fig. 4. Time series of annual average of the sea ice extent around Antarctica for observations for the 1987-2015 period and the control experiment of the models. The shading indicates the position of the two polynyas in the time series of the CM2.6 simulation only. The time series of observations has been shifted so that the concentration of atmospheric CO_2 in the model corresponds to the concentration of the first year of observations.

sign and magnitude to that of the observational dataset between years 15 and 60. Indeed, in the recovery period between the two polynyas, that is from the closing of the polynya until a return to the pre-polynya state, we observe a positive trend of 0.36 million km^2 per decade that compares well to 0.339 million km^2 for observations.

It is important to verify if this positive trend occurs only during the closing of the polynya or if it continues further. Indeed, as there is no polynya in the Southern Ocean for the period of observations analysed, this relationship would be meaningless if sea ice only increases during the period of closing. Using a series of maps of the sea ice concentration at the beginning of the Austral winter (July) for consecutive years, we ascertain that the first polynya closes at around year 30 of the simulation. Since the increasing trend in Fig. 4 continues until year 60, there is a period of approximately 30 years following the closing of the polynya that correspond to an increase in sea ice. It is unclear why that increase occurs and this constitutes a topic of further investigation. Considering that we have observed an open-ocean polynya in Antarctica from 1974-1976 (26), it is possible that the current positive trend in sea ice extent area could be due to a period of recovery from that polynya. However, this positive trend is found in the control experiment, which is a preindustrial simulation. When we look at the perturbation experiment, which includes anthropogenic forcing, we do not observe a positive trend but rather a stalling of the decreasing trend. This means that the positive sea ice trend induced by the polynya is compensated by the negative sea ice trend induced by anthropogenic forcing in the perturbation experiment. Still, these results suggest that the opening of the Weddell Sea polynya might have played an important role in the observed trend.

Role of Mesoscale Eddies in the Sea Ice Trend

In this section, we evaluate our second hypothesis that models with a higher ocean resolution, and consequently a better representation of mesoscale eddies, will simulate a greater response to anthropogenic forcing and as such a greater decrease in sea ice. To do so, we look at the response to climate change in the models. We subtract the control from the perturbation experiment to analyse purely the response to climate change. Then, we calculate the linear regression over time at each model grid cell for both sea ice concentration and SST (Fig.5). Overall sea ice concentration decreases (Fig.5 a.-c) while SST increases (Fig.5 d.-f.). We can see an area of intense SST warming along the western boundary current in all three models (Fig.5 d.-f.) and more warming along the Antarctic Circumpolar Current than close to Antarctica. CM2-1deg undergoes less of a decrease than CM2.5 or CM2.6 as indicated by the average decrease per decade, -0.5%/decade for CM2-1deg and -1%/decade or more for the other two. We can observe in CM2-1deg some areas of increase in sea ice corresponding to small areas of decrease in temperature although those trends are not significant. Overall, surface warming increases with resolution (0.12, 0.13 and 0.15 °C per decade for CM2-1deg (FIG. 5d.), CM2.5 (e.) and CM2.6 (f.) respectively). Hence, the model suite shows a clear link between a higher resolution on one hand and more warming and less sea ice on the other.

To evaluate if the models show evidence of the delayed warming mechanism as a response to climate change, we compute a Hovmöller diagram (latitude vs time) of the annual average of SST (Fig. 6). The first structure that we observe is a decrease in the SST for approximately the first two decades followed by a warming for the rest of the time period in all three models. This result supports Ferreira *et al.*(15) delayed warming mechanism (see Fig. 1). We note that the cooling phase is more prominent as the resolution increases. Indeed, CM2-1deg shows period of cooling within the warming period (Fig. 6a.). The second feature observed in Fig. 6 is that the warming is more intense as the resolution increases south of 50°S. This supports our previous results (see comments on Fig. 5).

A better representation of mesoscale eddies in models seems to be associated with a stronger warming of the SST in the Southern Ocean. These results suggest that mesoscale eddies have an important role in the response of Antarctic sea ice to climate change.

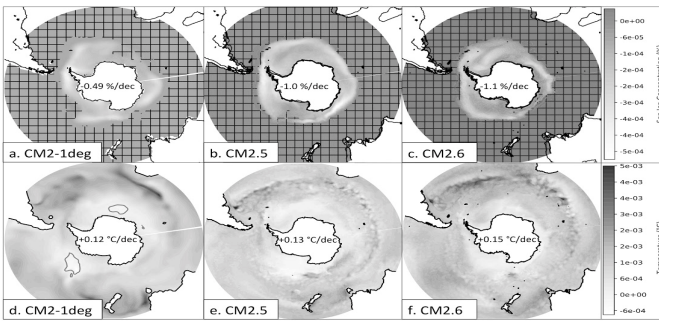


Fig. 5. Slope of the linear regression of the difference between the perturbation and control experiments of the sea ice concentration (a.-c.) and the sea surface temperature (d.-f.). We consider purely the response to climate change by subtracting the control from the perturbation experiment. The hashes point out trends that are not significant. Note the decadal trend (%/dec or °C/dec) inscribed on the Antarctic continent.

Discussion

The positive sea ice extent trend present in the control experiment of CM2.6 during the recovery period from the polynya is the only trend similar to that of observations detected in all models analysed (see Fig. 4). This trend in CM2.6 occurs during and after the closing of a Weddell Sea polynya, which is also the case in observations (1974-1976 polynya). As such, it supports our hypothesis that the recovery from the 1970's polynya can explain the current positive sea ice trend. This hypothesis could be further supported by the 2016 and 2017 sea ice trends. Indeed, we have observed the lowest sea ice extent seen in the satellite record in 2016 that was followed, in 2017, by the opening of a small Weddell Sea polynya.(10) If this polynya continues to grow in the next few years and the sea ice trend continues to decrease, it could offer supporting evidence that natural variability dominates over anthropogenic forcing for the sea ice trend in Antarctica.

Fig. 5 and 6 indicate a strong link between a higher ocean resolution and more intense warming in the Southern Ocean. Fig. 5 shows an increase in SST and a negative sea ice extent area trend for most of the Southern Ocean. The regionally averaged warming (30°S to 90°S) increases by 8% (CM2-1deg to CM2.5) to 15% (CM2.5 to CM2.6) with resolution. Bitz and Polvani (2012) (27) find the opposite result in their study which shows that the ocean warming observed due to atmospheric ozone loss is somewhat muted in their 0.1° ocean resolution model compared to their 1° ocean resolution model.(27) Differences between the models and experiments of Bitz and Polvani (2012) and ours are numerous and further investigation need to be done to elucidate the causes of the divergence in results. Fig. 6 also presents supporting evidence for the delayed warming mechanism of Ferreira *et al.*(15) caused by an increase in the westerly winds. This increase in the westerly winds has been observed in this suite of climate models (not shown). In the climate change scenario, we observe a cooling of the surface in the Southern Ocean followed by a strong warming that gets more intense as resolution increases. Our results supports the conclusions of Screen *et al.*(18) who demonstrated that models with a parameterization of eddies do not show

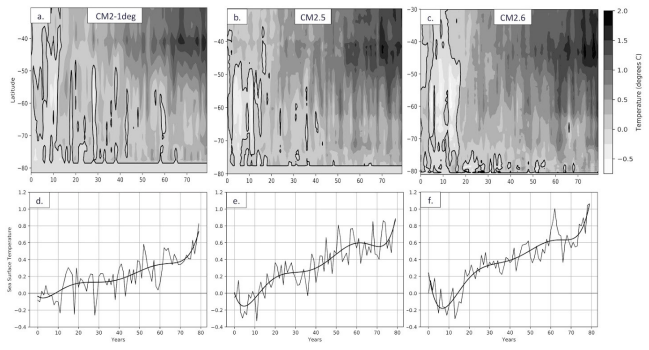


Fig. 6. a.-c. Hovmöller diagram (latitude vs time) of annual average of sea surface temperature for the difference between the sensitivity and control experiments (signal for anthropogenic climate change only). A solid line highlights the zero contours. d.-f. Annual average of the sea surface temperature in the Southern Ocean from 50°S to 90°S. A 6th degree polynomial fit is superimposed on the time series.

as strong of a warming as models that explicitly resolve mesoscale eddies. Both Fig. 5 and 6 support our second hypothesis that warming due to anthropogenic forcing will induce a decrease in sea ice that is augmented by the presence of mesoscale eddies.

If the current sea ice trend is indeed due to the recovery from the Weddell Sea polynya of the 1970s, it suggests a great influence of natural variability on the Antarctic sea ice cover, which might be underestimated in models like CM2-1deg or CM2.5. Because it resolves polynyas, CM2.6 might be more skilled to accurately simulate the Antarctic sea ice trend. However, an accurate simulation requires the model's natural variability be synchronized with that of the real world. It is unlikely though that such a synchronization will spontaneously occur in climate models.

Another source of misrepresentation of the sea ice trend in models relates to the response to anthropogenic forcing not being accurately simulated in models. One possibility is that natural variability could be masked in models by a strong response to anthropogenic forcing. If the natural and forced signals are opposed and if the anthropogenic forcing is too strong in models, then the trend would not be the same in models and in observations. Concurrently, the intensity of the long-term warming response might be too strong in this suite of models. Indeed, an increase of 1% per year in the perturbation experiment signifies that the doubling of the CO_2 concentration from preindustrial levels will occur over a period of 70 years only. In this intensity of the warming response, our results differ from those of Ferreira *et al.*(15) who use a weaker anthropogenic forcing that stay close to historical values. Indeed, they observe a maximum of 0.6°C of SST response in the ocean South of 30°S while we observe a maximum of 2.0°C. Still, the onset of the warming response is similar in time at around 20 years.

Conclusion

To conclude, our results clearly suggest the importance of representing polynyas to accurately simulate the Antarctic sea ice trend. Also, the role of mesoscale eddies in the response of Antarctic sea ice to climate change cannot be overlooked. Our results show their importance for the intensity of the warming response and subsequent decrease in sea ice present in models.

Acknowledgements

This work was supported by the Loretta del Bosco Science Undergraduate Research Award. We thank the Geophysical Fluid Dynamics Laboratory for producing and making available their model output. We thank the Atmospheric and Oceanic Sciences department at McGill University for providing office space and a stimulating environment for our research.

References

1. Armour KC, Bitz CM. Observed and projected trends in Antarctic sea ice. US CLIVAR Variations. 2015;13(4):12-9.
2. Cavalieri DJ, Parkinson CL. Antarctic sea ice variability and trends, 1979-20062008.
3. Turner J, Bracegirdle TJ, Phillips T, Marshall GJ, Hosking JS. An Initial Assessment of Antarctic Sea Ice Extent in the CMIP5 Models. Journal of Climate. 2013;26(5):1473-84.
4. Polvani LM, Smith KL. Can natural variability explain observed Antarctic sea ice trends? New modeling evidence from CMIP5. Geophysical Research Letters. 2013;40(12):3195-9.
5. Heuze C, Heywood KJ, Stevens DP, Ridley JK. Southern Ocean bottom water characteristics in CMIP5 models. Geophysical Research Letters. 2013;40(7):6.
6. Aguiar W, Mata MM, Kerr R. On deep convection events and Antarctic Bottom Water formation in ocean reanalysis products. Ocean Sci. 2017;13(6):851-72.
7. Morales Maqueda MA, Willmott AJ, Biggs NRT. Polynya Dynamics: a Review of Observations and Modeling. Reviews of Geophysics. 2004;42(1):n/a-n/a.
8. Zwally HJ. Antarctic sea ice, 1973-1976: satellite passive-microwave observations. Washington, D.C.: Scientific and Technical Information Branch, National Aeronautics and Space Administration : For sale by the Supt. of Docs., U.S. G.P.O.; 1983. xiv, 206 p. p.
9. Zanowski H, Hallberg R, Sarmiento JL. Abyssal Ocean Warming and Salinification after Weddell Polynyas in the GFDL CM2G Coupled Climate Model. Journal of Physical Oceanography. 2015;45(11):2755-72.
10. Stuecker MF, Bitz CM, Armour KC. Conditions leading to the unprecedented low Antarctic sea ice extent during the 2016 austral spring season. Geophysical Research Letters. 2017;44(17):9008-19.
11. Bintanja R, van Oldenborgh GJ, Drijfhout SS, Wouters B, Katsman CA. Important role for ocean warming and increased ice-shelf melt in Antarctic sea-ice expansion. Nat Geosci. 2013;6(5):376-9.
12. Swart NC, Fyfe JC. The influence of recent Antarctic ice sheet retreat on simulated sea ice area trends. Geophysical Research Letters. 2013;40(16):4328-32.
13. Holland PR, Kwok R. Wind-driven trends in Antarctic sea-ice drift. Nat Geosci. 2012;5(12):872-5.
14. Sigmond M, Fyfe JC. The Antarctic Sea Ice Response to the Ozone Hole in Climate Models. Journal of Climate. 2014;27(3):1336-42.
15. Ferreira D, Marshall J, Bitz CM, Solomon S, Plumb A. Antarctic Ocean and Sea Ice Response to Ozone Depletion: A Two-Time-Scale Problem. Journal of Climate. 2015;28(3):1206-26.
16. Thompson DWJ, Wallace JM. Annular modes in the extratropical circulation. Part I: Month-to-month variability. Journal of Climate. 2000;13(5):1000-16.
17. Armour KC, Marshall J, Scott JR, Donohoe A, Newsom ER. Southern Ocean warming delayed by circumpolar upwelling and equatorward transport. Nature Geosci. 2016;9(7):549-54.
18. Screen JA, Gillett NP, Stevens DP, Marshall GJ, Roscoe HK. The Role of Eddies in the Southern Ocean Temperature Response to the Southern Annular Mode. Journal of Climate. 2009;22(3):806-18.
19. Farneti R, Delworth TL, Rosati AJ, Griffies SM, Zeng F. The Role of Mesoscale Eddies in the Rectification of the Southern Ocean Response to Climate Change. Journal of Physical Oceanography. 2010;40(7):1539-57.
20. Meredith MP, Hogg AM. Circumpolar response of Southern Ocean eddy activity to a change in the Southern Annular Mode. Geophysical Research Letters. 2006;33(16):4.
21. Meier W, Fetterer F, Savoie M, Mallory S, Duerr R, Stroeve J. NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentratio Boulder, Colourado USA: National Snow and Ice Data Center; 2013 [updated 2016. Version 2:[
22. Staff NCfAR. The Climate Data Guide: SST data: HadiSST v1.1. 2017 [updated 22 Aug 2017. Available from: https://climatedataguide.ucar.edu/climate-data/sst-data-hadisst-v11.
23. Griffies SM, Winton M, Anderson WG, Benson R, Delworth TL, Dufour CO, et al. Impacts on Ocean Heat from Transient Mesoscale Eddies in a

Hierarchy of Climate Models. Journal of Climate. 2015;28(3):952-77.

24. Dufour CO, Morrison AK, Griffies SM, Frenger I, Zanowski H, Winton M. Preconditioning of the Weddell Sea Polynya by the Ocean Mesoscale and Dense Water Overflows. Journal of Climate. 2017;30(19):7719-37.

25. Cheon WG, Lee SK, Gordon AL, Liu Y, Cho CB, Park JJ. Replicating the 1970s' Weddell Polynya using a coupled ocean-sea ice model with reanalysis surface flux fields. Geophysical Research Letters. 2015;42(13):5411-8.

26. Carsey FD. MICROWAVE OBSERVATION OF THE WEDDELL POLYNYA. Mon Weather Rev. 1980;108(12):2032-44.

27. Bitz CM, Polvani LM. Antarctic climate response to stratospheric ozone depletion in a fine resolution ocean climate model. Geophysical Research Letters. 2012;39(20):n/a-n/a.

Submitted: 01/14/18

Review Article

[1]Department of Biology, McGill University, Montreal, QC, Canada

Keywords

Mercury, Methylation, Indicators, Aquatic System, DOC

Email Correspondence

sandrine.beaumont-courteau@mail.mcgill.ca

Sandrine Beaumont-Courteau[1]

Interactions between Mercury and Biogeochemical Features of Aquatic Ecosystems: A Critical Review

Abstract

Background: Mercury (Hg) is a pollutant known to affect the nervous system. The most threatening form of Hg is methylmercury (MeHg), which can biomagnify in aquatic biota. This critique summarizes our current understanding of the major relationships between mercury speciation in aquatic environments and various environmental factors to determine the most suitable indicators of MeHg pollution.

Methods: Searches were performed using Web of Science. Ultimately, 33 studies were chosen and reviewed. Priority was given to recent studies as this review focused on the current state of our knowledge.

Summary: The total amount of Hg (THg) alone cannot be used as an indicator of MeHg since the relationship between THg and MeHg is poorly correlated. Microbial DNA is often used in research to give insight into the mercury cycle and fate in aquatic systems, but further studies are needed to accurately assess MeHg concentration using DNA. Based on an early study, it was thought that water colour could indicate the amount of Hg in aquatic biota. However, subsequent work has shown this to not be the case since dissolved organic matter (DOM) can help or hinder Hg bioavailability and bioaccumulation. There is a nonlinear relationship between dissolved organic carbon (DOC) and MeHg bioaccumulation where there is a threshold concentration (~8.5 mg C L^{-1} DOC) above which mercury bioaccumulation is hindered. A recent study found that the methylation rate of an aquatic system was correlated to the type of organic matter present in the sediments, although it could not predict the amount of MeHg present in the water since runoff brings Hg from the catchment to the downstream aquatic system. Recent advancements have been made to understand components of the mercury cycle (e.g. land-water interactions, microbial methylation, water-sediment interactions), but a model encompassing all components has yet to be constructed.

Introduction

The Minamata Bay disaster that occurred in Japan in the 1950s raised awareness about the level of mercury (Hg) in aquatic environments.(1) Minamata Bay was contaminated with methylmercury (MeHg) discharged through the wastewater of a nearby chemical plant.(2) Many forms of Hg are hazardous neurotoxins to humans, but MeHg is the most concerning one since it can biomagnifiy throughout food webs.(2-4) By unknowingly ingesting contaminated fish and shellfish from the Minamata bay, many citizens suffered from methylmercury poisoning, also called "Minamata Disease".(2) Since Hg affects the nervous system, common symptoms of the Minamata Disease are visual, sensory and auditory disturbances as well as uncontrolled muscles movements.(2) A global interest to extensively study the Hg cycle was stimulated by this ecological disaster. This article presents an overview of our current knowledge on the Hg cycle in the aquatic ecosystem.

Global emission and deposition of Hg have increased significantly since the industrial revolution.(5,6) Anthropogenic sources of mercury emission include fuel combustion and industrial manufacturing.(4,7,8) As a global pollutant, the elemental form of mercury (Hg_0) can travel for tens of thousands of kilometers within the atmosphere before being deposited in an aquatic environment or before being retained within the soil or the vegetation.(8-10) Moreover, once deposited, Hg is not necessarily trapped. In fact, secondary Hg re-emission from its aquatic and terrestrial substrates is not negligible since it can contribute to 56-65% of the total Hg emitted. (10)

The Hg cycle is influenced by many biogeochemical factors that can make it difficult to predict its behaviour in different natural settings.(10,11) For example, in the atmosphere, Hg_0 is the main form of mercury, but within the aquatic and terrestrial systems, inorganic Hg (Hg(II)) is predominant. (12) In the aquatic systems, a small amount of Hg(II) is transformed into MeHg primarily through a biotic pathway, but an abiotic methylation process does exist.(7,11) Even if minute concentrations of MeHg are present within the water column, the concentration of Hg in the apex predators can be up to 10^7 times greater than that of the water.(1)

The present review outlines the current understanding of the major relationships between mercury speciation in aquatic environments and various environmental factors to find the best indicators of MeHg pollution. The relationship between the total amount of mercury (THg) and MeHg was considered with regards to the physical setting of the aquatic systems. The possibility of using microbial DNA in the sediments to detect Hg pollution was also examined, since the principal producers of MeHg are anaerobic microbes.(4) Moreover, the interactions between mercury species and organic matter (OM) were studied with a particular interest to lake browning, which is an increase in water colour caused by the increased export of coloured OM.(13)

Methylmercury

Relationship between THg and MeHg

Understanding the speciation of mercury is essential when assessing the impact of mercury on ecosystems or on human health. Since Hg(II) is the prevalent form of mercury in aquatic systems, the total amount of mercury (THg) within the water or in the sediments can be used as a proxy for the level of Hg(II) within these systems.(14,15) MeHg is produced from Hg(II), thus many studies have tried to examine the relationship between THg and MeHg.(1,16) Some studies found a positive relationship between THg and MeHg.(12,16) However, several other studies have found no consistent relationships between the concentration of THg and MeHg in both water and sediments, which suggests that THg concentration is not always reflective of MeHg concentration in various aquatic ecosystems.

(10,12,14,17,18)

The ratio of MeHg to THg was found to be inconsistent (10) and could only explain 25% of MeHg variations in model used by Fleck et al..(12) However, THg can be used to find some relationships. For example, Drott et al.(16) showed a significant relationship between the potential methylation rate constant (K_m) and the concentration of MeHg normalized to the total concentration of Hg. This relationship was not significant when only the bulk concentration of MeHg was used. Also, the ratio of MeHg to THg can be used to indicate the relative methylation efficiency of an ecosystem. (12) Thus, THg is a factor that contributes to the production of MeHg but, since the relationship is inconsistent, it should not be used alone as an indicator of MeHg concentration.(15,16)

MeHg and THg are also found in sediments.(14) However, Eagle-Smith et al.(10) found no relationship between the sediment THg and the fish THg and only a weak positive correlation was found between the concentration of MeHg in the sediments and the THg concentration in fish tissue. Overall, the results to date suggest that the amount of MeHg present in the sediment alone is not a good indicator of the MeHg cycling in the food web. Even though the relationship between THg and MeHg is not always reflective of the bioaccumulation of Hg in the food web, several pollution assessment methods use the THg concentration in sediments to determine the mercury pollution of aquatic systems.(10,12,19) Various sediment quality guidelines (SQG) were developed, which all have their own limitations and advantages.(20) In an effort to obtain values applicable to different aquatic systems, Macdonald et al.(20) gathered many SQG to evaluate them and obtained a consensus whereby the threshold effect concentration of Hg in a freshwater ecosystem was 0.18 mg/kg DW and the probable effect concentration was 1.06 mg/kg DW. However, this consensus was shown to be limited.

Gao et al.(19) determined that those SQG was not representative of China's aquatic ecosystems, since the results obtained with SQG did not match the results obtained by three other mercury pollution assessment methods: contamination factor, geoaccumulation index, and potential ecological risk. Furthermore, SQG are not universal since they ignore the bioavailability and the methylation rate of Hg in an aquatic system.(21) Conder et al.(21) suggest that SQG could be used only as an initial screening method.

Influence of landscape and environmental settings

In their study, Fleck et al.(12) demonstrated that MeHg concentration is more influenced by environmental conditions and by landscape features than by THg, since those characteristics accounted for 51% of the MeHg variation independently of the THg. The authors evaluated the THg and MeHg concentrations of various aquatic feature types (canals, estuaries, lakes and streams) and environmental settings (agriculture, forested, open-water, rangeland, wetland, urban) of western North America by using the data sets of several sources. Lakes and streams generally had the highest THg and MeHg concentrations. However, Fleck et al. (2016) also noticed that THg concentrations were highest in open-water, whereas MeHg concentrations were not. The methylation efficiency tends to be lower in estuaries, open-water, and urban settings (Fig. 1; 12).

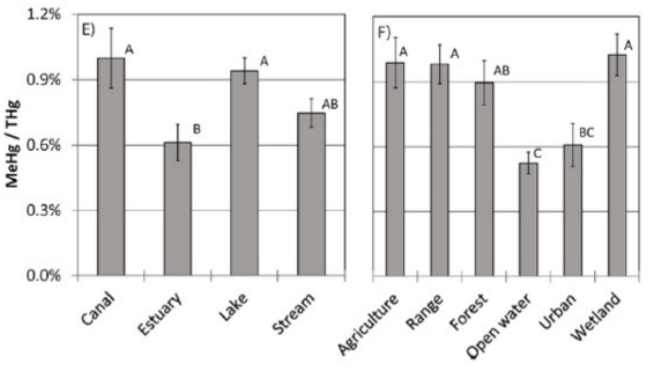


Fig. 1. Ratio of methylmercury (MeHg) to the total amount of mercury (THg). A) MeHg to THg ratio for the studied aquatic feature types, B) MeHg to THg ratio for the studied environmental settings. (Modified from Fleck et al., 2016)

Research on the Canadian Arctic freshwater systems showed that while rivers and streams tend to have higher concentrations of THg, it is in the ponds where MeHg concentration tends to be the highest.(22) In the Arctic, there is a tendency for ponds and wetlands to have a high MeHg concentration.(1,22) This might be due to the warm, shallow water that enhances bacterial activity.(22)

Furthermore, in recent years, Hg emissions in North America have decreased,(23) yet point sources still discharge Hg in the water systems, preventing these systems from responding to the decreased Hg emissions. (6) Drevnick et al.(6) found that for lakes not directly polluted by point sources, the time to respond to decreased emissions was inversely correlated with watershed size. Since the soil can sequester Hg and is a great Hg sink, the runoff water from the watershed carries Hg into the aquatic systems.(24) Larger watersheds will be slow to respond to the decreased Hg emission as the Hg moves slowly from the catchment to the aquatic environment.

Mercury methylation

Poor correlation between MeHg and THg suggests that the methylation of mercury does not depend solely on the quantity of inorganic mercury present within systems.(4,18) Since Hg(II) methylation is primarily microbial, the bioavailability of Hg(II) and the microbial productivity greatly influence the Hg methylation rate.(4,25) The main methylators are sulphate reducers, iron reducers and methanogens, which are present in surficial sediments, anoxic bottom waters and wetlands.(10,25) The ability to produce MeHg depends on the bacteria strain, not the genus. (4,26)

The capacity to methylate Hg is often found in the microbes possessing the gene cluster *hgc*A/*hgc*B.(27) Du et al.(26) studied the relationship between the abundance of certain genes in soils and sediments and the concentration of MeHg in those. *Dsr*B, a gene found in sulphate reducers, and *hgc*A, a gene for Hg methylation, were positively correlated with the concentration of MeHg, which suggests that bacteria with these genes contribute to the methylation of Hg.(26) In another study done by Poulain et al.(5), the mercuric reductase gene (*mer*A) was used to observe the bacterial response to the increase in Hg emissions. They concluded that *mer*A could be a potential tool to study the delivery of mercury to the aquatic systems, since the evolutionary response of microbes to changes in mercury deposition is fast and seems correlated to the changes in anthropogenic emissions.

Relationship Between Dissolved Organic Matter and Hg

Water Colour and Hg

Wescott and Kalff (28) established that water colour and pH could be used as indicators of zooplankton MeHg concentration, which can be used as a proxy for fish tissue MeHg concentration. However, the 24 lakes studied by Wescott and Kalff (28) did not allow for a large variation, which limited the applicability of their findings. Not all lakes respond to Hg additions in the same manner. Isidorova et al.(13) noted that lake browning increased the transport of Hg to the sediments, where the methylation of Hg would be pronounced.

Lake browning is often caused by an increase in dissolved organic matter (DOM) in the aquatic systems.(13,29) DOM is known to interact strongly with Hg and affect its cycling and fate in the aquatic systems.(30) There is often a correlation between the amount of organic matter (OM) and the THg present in the system,(30) as DOM is the main mediator of mercury into aquatic systems through catchment area.(31)

Kinetics of DOM and Hg interaction

French et al.(32) showed that dissolved organic carbon (DOC) influences the bioaccumulation of Hg in Arctic lakes. Their research identified the existence of a DOC threshold concentration at about 8.5 mg C L-1. Exceeding that, the DOC concentration starts to hinder the Hg bioaccumulation(Fig. 2; 32). In the study conducted by Isidorova et al.(13), the reference used to associate lake browning to the increase of Hg concentration in fish was research done by Hongve et al.(33)

Hongve et al.(33) studied the increase of total organic carbon (TOC) and the variation in fish tissue MeHg concentration in two lakes. The authors noticed an increase in fish Hg in the lake where the increase in TOC reached a concentration similar to the DOC threshold of 8.5 mg L^{-1} determined by French et al.(32), yet both studies do not agree perfectly. French et al.(32) observe a decrease in Hg bioaccumulation past the 8.5 mg L^{-1} threshold while Hongve et al.(33) observe and increase at 9 mg L^{-1}. However, Hongve et al.(33) measured the TOC, which is the combination of particulate organic carbon (POC) and DOC, in their studied lakes, while French et al.(32) only measured the DOC. The use of TOC instead of DOC might explain the subtle differences between the studies.

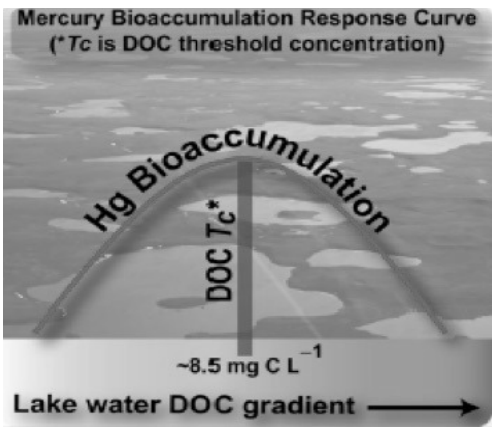


Fig. 2. Bell-shaped relationship between mercury (Hg) bioaccumulation and the concentration of dissolved organic carbon (DOC) present in the lake water. The DOC Tc represents the threshold after which DOC starts to hinder Hg bioaccumulation (Modified from French et al., 2014)

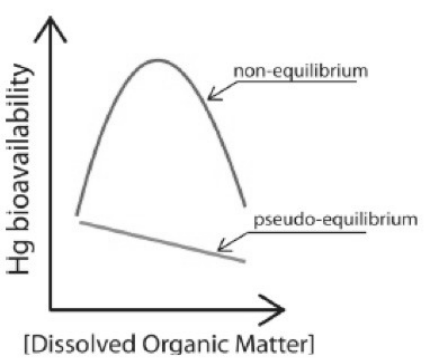


Fig. 3. Bell-shaped relationship between mercury (Hg) bioavailability and the concentration of dissolved organic matter (DOM) the water. The bell-shaped pattern is only present when Hg and DOM are in non-equilibrium condition. Hg bioavailability is lower and represented by a negative correlation in pseudo equilibrium conditions. (Modified from Chiasson-Gould et al., 2014)

Chiasson-Gould et al.(25) found that DOM was influenced by the bioavailability of Hg(II) to the bacteria and concluded that the presence of the same bell-shaped relationship between MeHg and the DOM suggests that the two variables are in non-equilibrium. During a short time period of less than 24 h, the freshly deposited Hg has more chance of being introduced in the food chain with the help of DOM (Fig. 3; 25).
Chiasson-Gould et al.(25) proposed a few hypotheses for this relationship. The first hypothesis is that when Hg first enters the aquatic system, it will bind to a small, kinetically accessible DOM. However, as time passes, Hg will bind to bigger, more stable DOM which will be too big to penetrate a bacterial cell wall.(25) The second hypothesis suggests that long exposure to DOM can change the bacterial cell wall properties.(25) The last hypothesis suggests the existence of a cycle of bioavailability and non-bioavailability.(25)

Collectively, these results suggest that browning may not be the best indicator of lake MeHg because the increase in water colour cannot always be associated with an increased of Hg in the aquatic biota since, after a certain threshold, DOM hinders Hg bioaccumulation. Recently, a study done by Bravo et al.(24) suggested that researchers should also focus on the type of DOM rather than only focusing on the concentration of DOM. Bravo et al.(24) found a relationship between the type of DOM and the methylation rate. The sediments containing more fresh algal, autochthonous organic matter tend to have a higher methylation rate than sediments dominated by terrigenous, allochthonous organic matter.(24) However, since the soil is a good sink for Hg and MeHg, there was a greater concentration of MeHg in the sediments of the terrigenous dominated lakes despite the lower methylation rate.(24) Further studies should be made on the relationships between the type of organic matter and the bioaccumulation of Hg.

Variables affecting MeHg	Relationships with MeHg	Suitability as a method to assess MeHg concentration
THg	• Poor correlation • THg/MeHg ratio varies for different physical setting • SQG are based on the THg concentration in the sediments	Suitable as a screening method only (SQG)
Microbial DNA	• Positive relationship between *dsr*B, *hgc*A and MeHg concentration • Consideration of *mer a* as a potential tool to study mercury delivery in the system	Potential method, but need further study
Water color	• Hg methylation hindered by DOC concentration above ~8.5 mg L^{-1}	Unsuitable
Types of DOM	• May influence the methylation rate of Hg • Can not predict the amount of MeHg in the water	Unsuitable at the moment, but very recent concept that need further study

Table 1. Summary of the relationships between methylmercury (MeHg) and various environmental parameters that are used to assess MeHg concentrations.

Conclusion

Mercury is a global pollutant that poses a threat to human health. Our understanding of the interactions between Hg and various biogeochemical factors is central to properly assess and control concentrations of MeHg. Although the relationship between THg and MeHg in the aquatic systems is inconsistent, THg is still a factor influencing MeHg concentration in the systems. The landscape and environmental settings also greatly influence the amount of MeHg present within a system. Furthermore, the relationship between DOM and Hg in the aquatic systems is complex since DOM can both hinder and facilitate Hg bioavailability and bioaccumulation. (25,32) Therefore, lake browning, which is caused by an increase of DOM in the water, is not reflective of the amount of Hg present in fish tissue. Recently, it was found that the type of OM is correlated with the methylation rate, although it does not indicate the amount of mercury present in the system.(24) Moreover, microbial DNA is becoming an important tool to understand Hg cycle and fate.

The Hg cycle is still not fully understood. Future research directions should focus on the relationship between the types of OM and the bioaccumulation of Hg. It would also be worthwhile to create a model that can integrate the multiple features studied in this review article (THg, land use, concentration and type of OM and the bacterial distribution) to assess

more efficiently mercury pollution in the aquatic systems.

Acknowledgements

I would like to thank my supervisor Professor Irene Gregory-Eaves and her lab members for all the help and the support.

References

1. Lehnherr I. Methylmercury biogeochemistry: a review with special reference to Arctic aquatic ecosystems. Environmental Reviews. 2014;22(3):229-43.
2. Harada M. Minamata Disease - Methylmercury Poisoning in Japan Caused by Environmental-Pollution. Critical Reviews in Toxicology. 1995;25(1):1-24.
3. Hall BD, Bodaly RA, Fudge RJP, Rudd JWM, Rosenberg DM. Food as the dominant pathway of methylmercury uptake by fish. Water Air and Soil Pollution. 1997;100(1-2):13-24.
4. Hsu-Kim H, Kucharzyk KH, Zhang T, Deshusses MA. Mechanisms Regulating Mercury Bioavailability for Methylating Microorganisms in the Aquatic Environment: A Critical Review. Environmental Science & Technology. 2013;47(6):2441-56.
5. Poulain AJ, Aris-Brosou S, Blais JM, Brazeau M, Keller W, Paterson AM. Microbial DNA records historical delivery of anthropogenic mercury. Isme Journal. 2015;9(12):2541-50.
6. Drevnick PE, Cooke CA, Barraza D, Blais JM, Coale KH, Cumming BF, et al. Spatiotemporal patterns of mercury accumulation in lake sediments of western North America. Science of the Total Environment. 2016;568:1157-70.
7. Rasmussen PE, Friske PWB, Azzaria LM, Garrett RG. Mercury in the Canadian environment: Current research challenges. Geoscience Canada. 1998;25(1):1-13.
8. Mohapatra SP, Nikolova I, Mitchell A. Managing mercury in the great lakes: An analytical review of abatement policies. Journal of Environmental Management. 2007;83(1):80-92.
9. Chalmers AT, Argue DM, Gay DA, Brigham ME, Schmitt CJ, Lorenz DL. Mercury trends in fish from rivers and lakes in the United States, 1969-2005. Environmental Monitoring and Assessment. 2011;175(1-4):175-91.
10. Eagles-Smith CA, Wiener JG, Eckley CS, Willacker JJ, Evers DC, Marvin-DiPasquale M, et al. Mercury in western North America: A synthesis of environmental contamination, fluxes, bioaccumulation, and risk to fish and wildlife. Science of the Total Environment. 2016;568:1213-26.
11. Ullrich SM, Tanton TW, Abdrashitova SA. Mercury in the aquatic environment: A review of factors affecting methylation. Critical Reviews in Environmental Science and Technology. 2001;31(3):241-93.
12. Fleck JA, Marvin-DiPasquale M, Eagles-Smith CA, Ackerman JT, Lutz MA, Tate M, et al. Mercury and methylmercury in aquatic sediment across western North America. Science of the Total Environment. 2016;568:727-38.
13. Isidorova A, Bravo AG, Riise G, Bouchet S, Bjorn E, Sobek S. The effect of lake browning and respiration mode on the burial and fate of carbon and mercury in the sediment of two boreal lakes. Journal of Geophysical Research-Biogeosciences. 2016;121(1):233-45.
14. Lessard CR, Poulain AJ, Ridal JJ, Blais JM. Steady-state mass balance model for mercury in the St. Lawrence River near Cornwall, Ontario, Canada. Environmental Pollution. 2013;174:229-35.
15. Cossa D, Garnier C, Buscail R, Elbaz-Poulichet F, Mikac N, Patel-Sorrentino N, et al. A Michaelis-Menten type equation for describing methylmercury dependence on inorganic mercury in aquatic sediments. Biogeochemistry. 2014;119(1-3):35-43.
16. Drott A, Lambertsson L, Bjorn E, Skyllberg U. Do potential methylation rates reflect accumulated methyl mercury in contaminated sediments? Environmental Science & Technology. 2008;42(1):153-8.
17. Kelly CA, Rudd JWM, Louis VL, Heyes A. Is Total Mercury Concentration a Good Predictor of Methyl Mercury Concentration in Aquatic Systems. Water Air and Soil Pollution. 1995;80(1-4):715-24.
18. Delongchamp TM, Ridal JJ, Lean DRS, Poissant L, Blais JM. Mercury transport between sediments and the overlying water of the St. Lawrence River area of concern near Cornwall, Ontario. Environmental Pollution. 2010;158(5):1487-93.
19. Gao B, Han LF, Hao H, Zhou HD. Pollution characteristics of mercury (Hg) in surface sediments of major basins, China. Ecological Indicators. 2016;67:577-85.
20. MacDonald DD, Ingersoll CG, Berger TA. Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems. Archives of Environmental Contamination and Toxicology. 2000;39(1):20-31.
21. Conder JM, Fuchsman PC, Grover MM, Magar VS, Henning MH. Critical Review of Mercury Sediment Quality Values for the Protection of Benthic Invertebrates. Environmental Toxicology and Chemistry. 2015;34(1):6-21.
22. Chetelat J, Amyot M, Arp P, Blais JM, Depew D, Emmerton CA, et al. Mercury in freshwater ecosystems of the Canadian Arctic: Recent advances on its cycling and fate. Science of the Total Environment. 2015;509:41-66.
23. Norton SA, Jacobson GL, Kopacek J, Navratil T. A comparative study of long-term Hg and Pb sediment archives. Environmental Chemistry. 2016;13(3):517-27.
24. Bravo AG, Bouchet S, Tolu J, Bjorn E, Mateos-Rivera A, Bertilsson S. Molecular composition of organic matter controls methylmercury formation in boreal lakes. Nature Communications. 2017;8.
25. Chiasson-Gould SA, Blais JM, Poulain AJ. Dissolved Organic Matter Kinetically Controls Mercury Bioavailability to Bacteria. Environmental Science & Technology. 2014;48(6):3153-61.
26. Du HX, Ma M, Sun T, Dai XZ, Yang CY, Luo F, et al. Mercury-methylating genes dsrB and hgcA in soils/sediments of the Three Gorges Reservoir. Environmental Science and Pollution Research. 2017;24(5):5001-11.
27. Gilmour CC, Podar M, Bullock AL, Graham AM, Brown SD, Somenahally AC, et al. Mercury Methylation by Novel Microorganisms from New Environments. Environmental Science & Technology. 2013;47(20):11810-20.
28. Westcott K, Kalff J. Environmental factors affecting methyl mercury accumulation in zooplankton. Canadian Journal of Fisheries and Aquatic Sciences. 1996;53(10):2221-8.
29. Forsberg C and Pertersen RC. A darkening of Swedish lakes due to increased humus inputs during the last 15 years, Interational Association of Theorical and Applied Limnology Proceedings,1990; 24(1), 289-292.
30. Ravichandran M. Interactions between mercury and dissolved organic matter - a review. Chemosphere. 2004;55(3):319-31.
31. Grigal, DF. Inputs and outputs of mercury from terrestrial watersheds: A review, Environmental Review,2002 ;10(1), 1–39.
32. French TD, Houben AJ, Desforges JPW, Kimpe LE, Kokelj SV, Poulain AJ, et al. Dissolved Organic Carbon Thresholds Affect Mercury Bioaccumulation in Arctic Lakes. Environmental Science & Technology. 2014;48(6):3162-8.
33. Hongve D, Haaland S, Riise G, Blakar I, Norton S. Decline of Acid Rain Enhances Mercury Concentrations in Fish. Environmental Science & Technology. 2012;46(5):2490-1.

Submitted: 1/14/18

Review Article

[1]Department of Biochemistyr, McGill University, Montreal, QC, Canada

Keywords

Phosphatase, PRL, CNNM, structural biology, cancer therapy

Email Correspondence

wenxuan.wang@mail.mcgill.ca

Wenxuan Wang[1]

Phosphatases of Regenerating Liver (PRL) as Therapeutic Targets in Cancer

Abstract

Background: Phosphatases of regenerating liver (PRL) represent a class of protein tyrosine phosphatases with oncogenic activity. PRL overexpression enhances cell proliferation, transformation, and promotes metastasis in many human cancers. Most notably, PRLs interact with a family of magnesium transporters, cyclin M (CNNM), to regulate intracellular Mg^{2+} levels. Thus, PRLs are attractive therapeutic targets given their role in oncogenic and tumour suppressor signaling pathways by modulating cellular growth.

Methods: Academic research articles were obtained by searching key terms in the PubMed database. This review specifically focuses on the articles that provided a comprehensive overview of PRLs, CNNMs, and small molecule inhibitors of PRLs.

Summary: This review discusses the role of PRLs in promoting cancer metastasis and explores current strategies for targeting PRL activity through the use of small molecule inhibitors. Although several potent PRL inhibitors have been discovered, improvements must be made prior to clinical applications. Therefore, understanding the molecular basis of PRL inhibition is essential for developing novel therapeutic agents in cancer treatments.

Introduction

Post-translational modification of proteins is a highly conserved process to install diverse functional groups onto synthesized proteins. One example of a reversible modification is phosphorylation, which involves the addition of a phosphate group onto the hydroxyl groups of a serine, threonine, or tyrosine residue.(1) In particular, protein tyrosine phosphorylation plays a fundamental role in physiological and pathological processes of eukaryotes.(2) Protein phosphorylation is implemented in regulating many signal transduction networks. Therefore, it is crucial to maintain a very precise balance of tyrosine phosphorylation and dephosphorylation in order to carry out proper cellular functions.(1) Two classes of enzymes exist to regulate the homeostasis of tyrosine phosphorylation: protein tyrosine kinases (PTKs), which catalyze tyrosine phosphorylation, and protein tyrosine phosphatases (PTPs), which are responsible for dephosphorylation.(1) PTPs are the largest family of phosphatases and are important modulators of signal transduction pathways.(2) PTPs are implicated in the control of cell growth, proliferation, differentiation, oncogenic transformation, and synaptic plasticity.(2) Disturbance in the balance between PTK and PTP activity results in abnormal tyrosine phosphorylation, which is linked to the development of many neoplastic diseases.(2) Phosphatases of regenerating liver (PRL) are a member of the PTP family that has implications in oncogenesis and cancer metastasis; thus, it is an attractive drug target for cancer therapies. The development of potent small molecule inhibitors to selectively target PRL activity suggests a promising approach for the treatment of many human cancers.(2)

PRL Phosphatases

PTPs are integral components of many signaling transduction cascades, such as the MAP kinase pathway.(2) The PTP family includes 107 members that can be subdivided into four classes.(3) Class I, II, and III PTPs utilize cysteine-based catalysis, whereas class IV utilizes aspartate-based catalysis.(3) Class I constitutes the largest and most diverse class, which includes tyrosine-specific classical PTPs containing receptor-like and non-transmembrane PTPs, and dual specificity phosphatases (DSPs).(3) DSPs are unique from the classical PTPs due to their ability to dephosphorylate both tyrosine and serine/threonine residues. Thus, DSPs tend to have the most diverse substrate specificity.(3) PRLs are classified as DSPs within the class I PTP family.(3) PRLs demonstrate oncogenic activity in which their overexpression promotes cell proliferation, transformation, and metastasis in several human cancers.(3) The PRL family is composed of three closely related members: PRL1, PRL2, and PRL3. PRL1 was the first member identified, and was initially characterized as a strongly upregulated immediate-early gene in the regenerating liver following partial hepatectomies.(4) All PRL proteins share high amino acid sequence similarity; PRL1 and PRL2 exhibit 87% similarity, while PRL3 exhibits 76% and 79% similarity to PRL1 and PRL2, respectively.(5) PRLs have low sequence homology with other DSPs (normally less than 30%), with the closest structural homology to VHR, Cdc14, and PTEN dual specificity phosphatases.(5)

Expression and Cellular Localization of PRLs

Although PRL family members possess similar sequences, the distribution of PRL proteins varies in eukaryotic cells. C.M. Dumanual et al. conducted an extensive analysis of PRL expression in human tissues. Among the three PRL phosphatases, PRL2 expression is the most abundant and is ubiquitously expressed in almost all tissues, while PRL1 appears to be more varied between tissue types.(6) Indeed, the widespread nature of PRL1 and PRL2 expression suggests their basic functions are common to many cell types. In contrast, PRL3 has a much more restricted expression pattern; it is primarily found in the heart and skeletal muscles, and is generally expressed at a lower level than PRL1 and PRL2.(6) Tissue specificity of PRL expression suggests non-overlapping functions for different PRLs. The highly controlled expression of PRL3 may explain why its overexpression is associated with the development of metastatic carcinomas.(7) In contrast, in mice, PRL1 and PRL3 expression levels are much more restricted, showing expression in the colon and intestine, but have very low or no expression in other organ systems.(8)

A unique feature of PRLs not found in other PTPs is the presence of a CAAX prenylation motif in the C-terminal domain preceded by a polybasic region.(9) The prenylation motif facilitates intracellular localization of proteins to the plasma membrane.(9) Furthermore, PRLs are associated with the plasma membrane and early endosomes in mammalian cells.(9) PRL proteins with a mutated or deleted CAAX sequence re-localize to the cytosolic or nuclear fraction.(9) The polybasic region adjacent to the prenylation site also promotes membrane localization of the PRLs, where complete substitution of six basic residues with alanine abolishes the plasma membrane association of PRL1. Additionally, membrane localization is likely controlled by the cell cycle, as all three PRLs have been suggested

to regulate cell cycle progression in mitosis.(10) This suggests that cellular localization of PRLs to the plasma membrane is highly correlated with their function in tumour metastasis.

PRL Structural Features

PRLs are approximately 20 kDa and are one of the smallest PTPs that consist of a single catalytic domain and lack a regulatory domain.(11) PRLs contain two sequence motifs central to the catalytic mechanisms utilized by the PTP family: the signature active-site motif $HC(X)_5R$ in the P-loop and the WPD loop motif.(12) The catalytic cysteine in the phosphatase motif carries out a nucleophilic attack upon phosphorylated substrates, resulting in the formation of a thiophosphoryl enzyme intermediate.(11) The microenvironment within the active site of PTPs causes the catalytic cysteine residue to have a particularly low pKa, around 5.(11) This allows the cysteine side chain to exist as a thiolate at physiological pH and to act as a nucleophile.(11) The essential role of the catalytic cysteine was confirmed by the C104A mutant, in which loss of the cysteine thiolate abolishes the catalytic activity of PRL3.(11) The P-loop also contains a conserved arginine involved in the stabilization of the transition state by providing a positive charge.(12) Correct positioning of this flexible P-loop was found to be critical for the activation and enzymatic catalysis in many PTPs.(12) Moreover, the neighbouring WPD loop in the active site plays a functional role in catalysis, as it contains a conserved aspartic acid which participates as the proton donor in both the formation and hydrolysis of the phosphoenzyme intermediate.(12)

PRL Phosphatase Activity

As a member of the PTP family, PRL activity occurs through a two-step catalytic mechanism, which involves the formation of a phosphoenzyme intermediate (Fig. 1). Previous studies have found low catalytic activity with PRLs. Specifically, wild-type PRL3 activity was three orders of magnitude lower than that of a typical DSP, such as CDC25.(13) Several differences between PRLs and classical PTPs may account for their low catalytic activity. Most notably, PRLs contain an alanine instead of the highly conserved serine/tyrosine residue next to the invariant arginine in the phosphatase motif.(11) The serine/tyrosine hydroxyl group normally functions in the breakdown of the phosphocysteine intermediate.(11) Consequently, this substitution in PRLs results in initial burst kinetics followed by slow turnover of the phosphoenzyme intermediate.(11) Mutagenesis studies confirm this catalytic mechanism, as the substitution of alanine to serine significantly improved both the burst kinetic rate and catalytic efficiency.(11) Similarly, the rate-limiting step of the catalytic mechanism has been identified as the hydrolysis of the phosphoenzyme intermediate, as it is extremely long-lived, with a half-life of over one hour.(11) G. Kozlov et al. show that this is partially responsible for the very low catalytic activity of PRL3, and suggests that the missing hydroxyl group may be provided by the substrate.(11) However, the physiological substrates of PRLs are still largely unknown due to the slow overall rate of catalysis.(11)

Regulation of PRL through Oxidation

The conservation of C49 in the active site of PRL phosphatases suggests that it has a functional role. There is growing interest in the redox of the catalytic cysteine in PTPs and its role in the regulation of signaling pathways in response to oxidative stress.(15) Recent data indicates two possible mechanisms for the oxidation of the catalytic cysteine side chain that involve its conversion to a sulfonic acid or the formation of an intramolecular disulfide bond.(16) PRL1 and PRL3 are capable of forming an intramolecular disulfide between C49 and the catalytic cysteine C104.(17) In PRL2, this bond forms at analogous positions between C46 and C101.(17) The formation of the redox-dependent disulfide bond results in the loss of the catalytic cysteine thiolate, which blocks substrate binding and catalysis.(11) Since PRLs are prone to oxidation, this feature may also contribute to the low dephosphorylation activity observed.(11) Studies have shown that PRL1 can be oxidized *in vivo* by H_2O_2 treatment, where its redox status is controlled by the cellular glutathione system.(17) Several other DSPs are also subject to redox regulation, such as PTEN, CDC25, and MKP formation of disulfide bonds following H_2O_2 treatment.(18) Thus, the disulfide bond between C49 and catalytic cysteine C104 in PRLs is thought to protect the catalytic cysteine from irreversible oxidation during oxidative stress.(17)

PRL Interaction with Cyclin M (CNNM) Magnesium Transporters

PRL have been identified as oncogenes, yet no physiological substrate has been identified due to their slow catalytic rate.(11) However, recent studies demonstrate PRLs can bind to cyclin M (CNNM) family proteins.(19) This interaction is independent of their phosphatase activity and indicates a potential oncogenic mechanism through magnesium homeostasis regulation.(19)

The CNNM family proteins, also known as ancient conserved domain proteins, are transmembrane proteins involved in mediating Mg^{2+} efflux in mammalian cells.(20) The CNNM family is composed of four members, CNNM 1-4.(20) CNNM2 and CNNM3 are ubiquitously expressed in mouse tissue, CNNM1 is mainly in the brain, and CNNM4 is within the gastrointestinal tract.(21) CNNMs demonstrate significant sequence homology to other Mg^{2+} membrane transporters.(22) CNNMs contain an extracellular region, transmembrane domain, conserved cystathionine-B synthase (CBS) pair domain, and putative cyclic nucleotide binding domain.(22) In particular, the CBS-pair domain is conserved within other bacterial Mg^{2+} membrane transporters, such as CorC and MgtE.(22) The CNNM CBS-pair domain forms a dimer and is generally involved in nucleotide binding, particularly to adenylate nucleotides.(22) The Mg^{2+}-ATP complex induces a conformation change in CBS-pair domain dimers to regulate protein function.(22) CNNM transporters play a significant role in maintaining intracellular Mg^{2+} homeostasis for proper cellular functions.(23) Mg^{2+} is the most abundant divalent cation and is required as a cofactor for many enzymes involved in energy metabolism and genomic stability.(23) Moreover, increased intracellular Mg^{2+} is associated with increased cell proliferation.(23) CNNM-dependent Mg^{2+} efflux suppresses tumour progression by regulating energy metabolism and AMPK/mTOR signaling.(23)

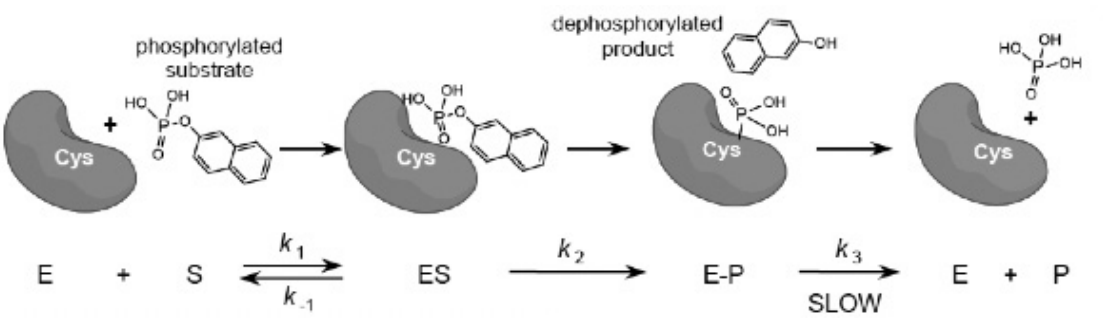


Fig. 1. Catalytic cycle of PRL phosphatases. PRL catalytic mechanisms occur through two steps. Initial burst kinetics is followed by a long-lived phosphocysteine intermediate. Hydrolysis of the intermediate is the rate-limiting step. (Fig. from Gulerez *et al.*(14))

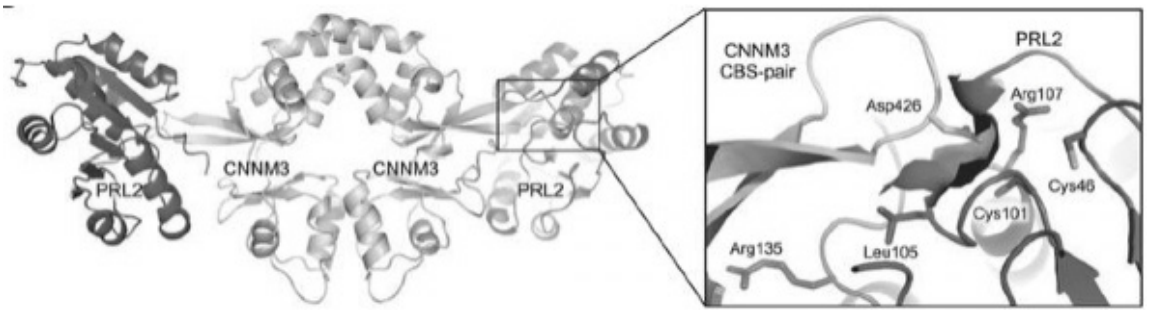


Fig. 2. Complex of PRL2 with the CBS-pair domain of CNNM3. CNNMs form a central homodimer that binds PRLs via an extended loop. A CNNM aspartic acid residue inserts into the phosphatase active site and is essential for complex formation. The catalytic cysteine contacts the CNNM aspartic acid to close the substrate pocket. (Fig. from Gulerez *et al.*(14))

PRLs interact with CNNM transporters to regulate Mg^{2+} transport.(14) CNNM and PRL complex formation is mediated by an extended loop in the CBS-pair domain, which contacts the PRL active site.(14) Mutagenesis studies by Gulerez et al. demonstrate a key D426 residue in CNNM3 that plays an important role in binding (Fig. 2).(14) The aspartate residue is inserted into the catalytic pocket and likely mimics the negatively charged phosphate group of a bound substrate. Therefore, PRLs act as pseudo-phosphatases, which are specific for CNNM proteins.(14) Additionally, the PRL-CNNM interaction is negatively regulated by the phosphorylation and oxidative state of PRL catalytic cysteine.(14) The phosphocysteine intermediate blocks CNNM binding through steric and electrostatic repulsion of D426.(14) The strong association between PRLs and metastatic diseases may be explained by its interaction with CNNM tumour suppressors to disrupt Mg^{2+} homeostasis.(23) PRL phosphatases are endogenously cysteine-phosphorylated in response to intracellular Mg^{2+} levels.(14) Through direct interactions, PRL binding inhibits CNNM-dependent Mg^{2+} efflux to raise intracellular Mg^{2+} levels to promote proliferation.(14) This supports the observation that Mg^{2+} deprivation leads to decreased PRL phosphorylation and increased total PRL levels, which suggests that the oncogenic property of PRLs is likely dependent on its interaction with CNNM Mg^{2+} transporters.

CNNM-PRL interaction is also regulated by PRL active site oxidation. (14) Oxidation of the PRL catalytic cysteine leads to decreased CNNM CBS-pair domain binding affinity.(14) Oxidation of the catalytic cysteine to form a disulfide bond induces an active site conformation change that inhibits CNNM binding (Fig. 3).(14) Catalytic cysteine phosphorylation may also be controlled by oxidation of the catalytic site, as PRL phosphorylation can only occur in the fully reduced form.(14)

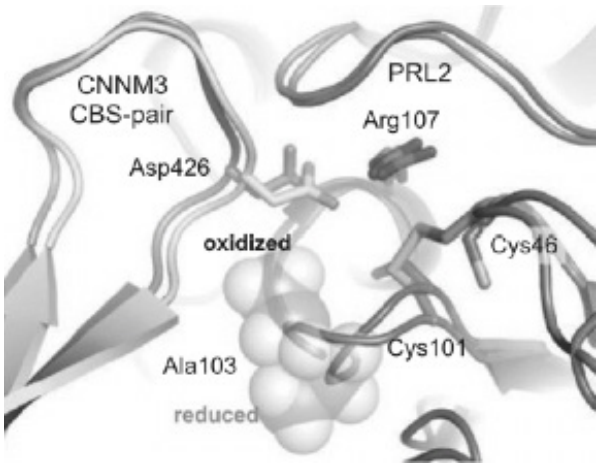


Fig. 3. Structural change of PRLs upon oxidation. A conformational change can be observed between the reduced and oxidized PRL2 complexes. Upon oxidation, the catalytic cysteine, Cys101, moves away from the CNNM3 Asp426 and Ala103 flips to displace CNNM3. (Fig. from Gulerez *et al.*(14))

Role of PRLs in Cancer

PRLs were first identified as a potential oncogene in 2001 through gene expression profiling.(7) S. Saha et al. discovered that PRL3 mRNA was consistently elevated in all metastatic lesions derived from colourectal cancer, whereas minimal PRL expression was observed in non-metastatic samples.(7) Aberrant elevation of PRL3 expression was also found in breast, lung, cervical, ovarian, and gastric cancers.(7) PRL3 expression is correlated with disease progression in ovarian cancers, with elevated levels in advanced stages.(24) Finally, increased tumour invasiveness is commonly observed in breast cancer patients expressing high levels of PRL3, suggesting that PRL3 expression can be used as a prognostic factor to predict advanced stages of disease.(25)

Enhanced PRL expression in cells results in increased cell adhesion, migration, invasiveness, and proliferation. Q. Zeng et al. demonstrated that PRL1 and PRL3 overexpression promotes cell mobility, cell invasiveness, and metastasis in Chinese hamster ovary cells, whereby catalytically inactive PRL3 reduced the ability to promote migration.(26) Y. Wang et al. also revealed that PRL2 can affect cell migration and invasion with human lung cancer cells.(27) PRL2 knockdown by short hairpin RNA significantly reduced tumour cell migration and invasion.(27) Both studies indicate that the oncogenic effect of PRLs is dependent on its phosphatase activity, as PRLs with mutated catalytic cysteine and arginine lose the ability to promote cell migration and metastasis.(27)

Furthermore, PRLs have been shown to affect cell apoptosis and angiogenesis. Recent studies reveal a novel connection between PRLs and p53 tumour suppressor.(28) S. Basak et al. demonstrated that, upon DNA damage, upregulation of PRL3 occurred in a p53-dependent manner to induce cell cycle arrest.(28) This mechanism occurs through increased Akt activation, negative feedback of the PI3K/Akt pathway, and transcription of growth arrest genes.(28) Similarly, p19Arf was upregulated in PRL3 knockout cells leading to MDM2 sequestration and p53-dependent cell cycle arrest.(28) PRL3 has also been found to facilitate angiogenesis, as it is strongly expressed in tumour vasculature.(29) PRL3 is involved in triggering tumour angiogenesis by downregulation of interleukin-4 to attenuate its inhibitory effect on vasculature formation.(29)

PRL Regulation of Signaling Pathways

PRL3 has been shown to alter several major oncogenic and tumour suppressor cell signal transduction pathways, including PTEN, p53, and Src pathways.(30) Notably, PRL3 appears to activate Src and PI3K/Akt signaling by reducing the expression of a negative regulator, Csk and PTEN respectively, to promote cellular growth, proliferation, and survival.(30) F. Liang et al. showed that PRL3 overexpression in HEK293 cells results in Csk downregulation and increased Src kinase activity.(31) Src pathway activation increases phosphorylation of downstream targets ERK1/2, STAT3, and p130CAS, leading to increased cell mobility and growth.(31) Specifically, ERK1/2 kinase activation contributes to a sustained G1 to S phase of cell cycle progression and proliferation.(31)

PRL interaction with adhesive proteins, such as cadherin and integrin, are also involved in modulating cell migration and invasiveness.(32) PRLs can interact with integrin proteins to enhance the binding to Src kinases to activate downstream signaling pathways.(33) Furthermore, E-cadherin and vinculin are downregulated with PRL3 overexpression, resulting in PI3K/Akt pathway activation by PTEN downregulation, a negative regulator of PI3K.(32) PRL-mediated signaling can also be explained by the activation of an extensive signaling network via receptor tyrosine kinases (RTKs). PRL3 regulates epidermal growth factor receptor (EGFR) by transcriptionally downregulating PTP1b, resulting in EGFR hyperphosphorylation and activation, which promotes growth.(34) Altogether, these studies strongly suggest that PRL3 overexpression is linked to the development of metastatic cancers by activating signaling pathways to enhance cellular proliferation. Therefore, PRL3 makes a very attractive target for small molecule inhibitors to halt the advancement of tumour progression.

Targeting PRLs using Small Molecule Inhibitors

Given that PRL overexpression is strongly associated with tumour progression and metastasis, there is increasing interest in discovering novel therapeutic agents to target the oncogenic properties of PRLs. Several small molecule inhibitors have been reported to inhibit PRL activity. For example, pentamidine was discovered to inhibit the activity of several PTPs, including PRL *in vitro*.(35) Pentamidine treatment is effective at inhibiting tumour growth induced by human melanoma cells.(35) However, pentamidine demonstrates nonspecific inhibition of all three PRLs, as well as an inhibitory effect on several other PTPs, such as PTP1b and MAPK phosphatase. Therefore, it cannot be concluded that the effect on tumour growth was solely due to PRL inhibition.(35)

More recently, high-throughput screening of the Roche chemical library revealed rhodanine derivatives and thienopyridone inhibit PRL3 activity.(36) Specifically, benzylidene rhodanine derivatives demonstrate greater potency of PRL3 inhibition than pentamidine, and is also more effective in decreasing the invasiveness of mouse melanoma cells.(36) The most noteworthy compound identified from the Roche chemical library was thienopyridone (7-amino-2-phenyl-5H-thieno-[3,2-c]pyridine-4-one), which exhibits an IC_{50} value of 132 nM against PRL3.(37) It displays a very high selectivity towards PRLs with minimal activity against 11 other phosphatase *in vitro*.(37) Furthermore, thienopyridone significantly inhibits tumour cell anchorage dependent growth of colon cancer cells and suppresses cell migration through p130Cas cleavage induction.(37) However, the main concern with thienopyridone is its high electron density, which may potentially cause idiosyncratic drug toxicity.(38) An improved inhibitor, iminothienopyridinedione (7-iminothieno-[3,2-c]pyridine-4,6(5H,7H)-dione), can be derived from thienopyridone through photooxygenation (Fig. 4).(38) It exhibits a greatly improved IC_{50} of 18 nM against PRL3, making it the most potent PRL3 inhibitor reported to date.(37) One advantage of iminothienopyridinedione is its decreased electron density and potential redox liability.(38) It is also 10-fold more potent than thienopyridone and has greater stability in solution.(38) These properties are important in decreasing off-target toxicity for improved drug absorbance and pharmacological response.(38)

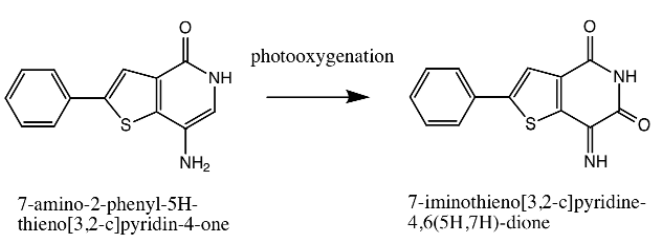


Fig. 4. Iminothienopyridinedione produced from thienopyridone through photooxygenation. The improved inhibitor iminothienopyridinedione has a 10-fold potency compared to thienopyridone. (Fig. from Salamoun *et al.*(38))

Conclusion

Phosphatases of regenerating liver are protein tyrosine phosphatases that play a critical role in cancer progression and tumour metastasis by inducing cell proliferation, survival, migration, and invasion. PRLs associate with CNNM magnesium transporters to promote oncogenic transformation through the regulation of intracellular magnesium levels. Evidence indicates that PRLs affect many important oncogenic and tumourigenic pathways, thus making it an attractive therapeutic target for cancer treatments. Several potent and selective PRL inhibitors, such as thienopyridone, have been discovered as promising anti-cancer agents. Although it is known that the oncogenic property of PRLs is dependent on its phosphatase activity, the mechanism in which small molecule inhibitors decrease PRL activity is still unknown and remains a compelling area of research. Determining the molecular basis of PRL inhibition is essential for the development of highly selective PRL inhibitors with improved pharmacological properties.

Acknowledgements

I would like to thank Dr. Kalle Gehring for having me in his laboratory. He not only provided me guidance and support in my work, but also in my future aspirations. It was an amazing opportunity to work on PRL-CNNMs and the experience I have gained will be incredibly valuable for pursuing research in the future. I would also like to thank my supervisor Dr. Guennadi Kozlov for the tremendous support he provided in the planning and analyzing of experiments. His mentorship and advice has been essential throughout the entire project.

References

1. Hunter T. Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. Cell. 1995;80(2):225-36.
2. Tonks NK. Protein tyrosine phosphatases: from genes, to function, to disease. Nature reviews Molecular cell biology. 2006;7(11):833-46.
3. Alonso A, Sasin J, Bottini N, Friedberg I, Friedberg I, Osterman A, et al. Protein tyrosine phosphatases in the human genome. Cell. 2004;117(6):699-711.
4. Mohn KL, Laz T, Hsu J, Melby A, Bravo R, Taub R. The immediate-early growth response in regenerating liver and insulin-stimulated H-35 cells: comparison with serum-stimulated 3T3 cells and identification of 41 novel immediate-early genes. Molecular and Cellular Biology. 1991;11(1):381-90.
5. Campbell AM, Zhang Z-Y. Phosphatase of regenerating liver: a novel target for cancer therapy. Expert opinion on therapeutic targets. 2014;18(5):555-69.
6. Dumaual CM, Sandusky GE, Crowell PL, Randall SK. Cellular localization of PRL-1 and PRL-2 gene expression in normal adult human tissues. Journal of Histochemistry & Cytochemistry. 2006;54(12):1401-12.
7. Saha S, Bardelli A, Buckhaults P, Velculescu VE, Rago C, Croix BS, et al. A phosphatase associated with metastasis of colourectal cancer. Science. 2001;294(5545):1343-6
8. Li J, Guo K, Koh VWC, Tang JP, Gan BQ, Shi H, et al. Generation of PRL-3-and PRL-1-specific monoclonal antibodies as potential diagnostic markers for cancer metastases. Clinical Cancer Research. 2005;11(6):2195-204.
9. Zeng Q, Si X, Horstmann H, Xu Y, Hong W, Pallen CJ. Prenylation-dependent association of protein-tyrosine phosphatases PRL-1,-2, and-3 with the plasma membrane and the early endosome. Journal of Biological Chemistry. 2000;275(28):21444-52.
10. Wang J, Kirby CE, Herbst R. The tyrosine phosphatase PRL-1 localizes to the endoplasmic reticulum and the mitotic spindle and is required for normal mitosis. Journal of Biological Chemistry. 2002;277(48):46659-68.
11. Kozlov G, Cheng J, Ziomek E, Banville D, Gehring K, Ekiel I. Structural insights into molecular function of the metastasis-associated phosphatase PRL-3. Journal of Biological Chemistry. 2004;279(12):11882-9.
12. Sun J-P, Wang W-Q, Yang H, Liu S, Liang F, Fedorov AA, et al. Structure and biochemical properties of PRL-1, a phosphatase implicated in cell growth, differentiation, and tumor invasion. Biochemistry. 2005;44(36):12009-21.
13. Gottlin EB, Xu X, Epstein DM, Burke SP, Eckstein JW, Ballou DP, et al. Kinetic analysis of the catalytic domain of human cdc25B. Journal of Biological Chemistry. 1996;271(44):2 7445-9.
14. Gulerez I, Funato Y, Wu H, Yang M, Kozlov G, Miki H, et al. Phosphocysteine in the PRL-CNNM pathway mediates magnesium homeostasis. EMBO reports. 2016;17(12):1890-900.

15. Meng T-C, Fukada T, Tonks NK. Reversible oxidation and inactivation of protein tyrosine phosphatases in vivo. Molecular cell. 2002;9(2):387-99.

16. Salmeen A, Andersen JN, Myers MP, Meng T-C, Hinks JA, Tonks NK, et al. Redox regulation of protein tyrosine phosphatase 1B involves a sulphenyl-amide intermediate. Nature. 2003;423(6941):769-73.

17. Yu L, Kelly U, Ebright JN, Malek G, Saloupis P, Rickman DW, et al. Oxidative stress-induced expression and modulation of Phosphatase of Regenerating Liver-1 (PRL-1) in mammalian retina. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research. 2007;1773(9):1473-82.

18. Cho S-H, Lee C-H, Ahn Y, Kim H, Kim H, Ahn C-Y, et al. Redox regulation of PTEN and protein tyrosine phosphatases in H2O2-mediated cell signaling. FEBS letters. 2004;560(1-3):7-13.

19. Hardy S, Uetani N, Wong N, Kostantin E, Labbe D, Bégin L, et al. The protein tyrosine phosphatase PRL-2 interacts with the magnesium transporter CNNM3 to promote oncogenesis. Oncogene. 2015;34(8):986-95.

20. Yamazaki D, Funato Y, Miura J, Sato S, Toyosawa S, Furutani K, et al. Basolateral Mg2+ extrusion via CNNM4 mediates transcellular Mg2+ transport across epithelia: a mouse model. PLoS genetics. 2013;9(12):e1003983.

21. de Baaij JH, Stuiver M, Meij IC, Lainez S, Kopplin K, Venselaar H, et al. Membrane topology and intracellular processing of cyclin M2 (CNNM2). Journal of Biological Chemistry. 2012;287(17):13644-55.

22. Gómez-García I, Stuiver M, Ereño J, Oyenarte I, Corral-Rodríguez MA, Müller D, et al. Purification, crystallization and preliminary crystallographic analysis of the CBS-domain pair of cyclin M2 (CNNM2). Acta Crystallographica Section F: Structural Biology and Crystallization Communications. 2012;68(10):1198-203.

23. Funato Y, Yamazaki D, Mizukami S, Du L, Kikuchi K, Miki H. Membrane protein CNNM4–dependent Mg2+ efflux suppresses tumor progression. The Journal of clinical investigation. 2014;124(12):5398.

24. Polato F, Codegoni A, Fruscio R, Perego P, Mangioni C, Saha S, et al. PRL-3 phosphatase is implicated in ovarian cancer growth. Clinical Cancer Research. 2005;11(19):6835-9.

25. Radke I, Götte M, Kersting C, Mattsson B, Kiesel L, Wülfing P. Expression and prognostic impact of the protein tyrosine phosphatases PRL-1, PRL-2, and PRL-3 in breast cancer. British journal of cancer. 2006;95(3):347-54.

26. Zeng Q, Dong J-M, Guo K, Li J, Tan H-X, Koh V, et al. PRL-3 and PRL-1 promote cell migration, invasion, and metastasis. Cancer research. 2003;63(11):2716-22.

27. Wang Y, Lazo JS. Metastasis-associated phosphatase PRL-2 regulates tumor cell migration and invasion. Oncogene. 2012;31(7):818-27.

28. Basak S, Jacobs SB, Krieg AJ, Pathak N, Zeng Q, Kaldis P, et al. The metastasis-associated gene Prl-3 is a p53 target involved in cell-cycle regulation. Molecular cell. 2008;30(3):303-14.

29. Guo K, Li J, Wang H, Osato M, Tang JP, Quah SY, et al. PRL-3 initiates tumor angiogenesis by recruiting endothelial cells in vitro and in vivo. Cancer research. 2006;66(19):9625-35.

30. Bessette DC, Qiu D, Pallen CJ. PRL PTPs: mediators and markers of cancer progression. Cancer and Metastasis Reviews. 2008;27(2):231-52.

31. Liang F, Liang J, Wang W-Q, Sun J-P, Udho E, Zhang Z-Y. PRL3 promotes cell invasion and proliferation by down-regulation of Csk leading to Src activation. Journal of Biological Chemistry. 2007;282(8):5413-9.

32. Wang H, Quah SY, Dong JM, Manser E, Tang JP, Zeng Q. PRL-3 down-regulates PTEN expression and signals through PI3K to promote epithelial-mesenchymal transition. Cancer research. 2007;67(7):2922-6.

33. Peng L, Jin G, Wang L, Guo J, Meng L, Shou C. Identification of integrin α1 as an interacting protein of protein tyrosine phosphatase PRL-3. Biochemical and biophysical research communications. 2006;342(1):179-83.

34. Al-aidaroos AQO, Yuen HF, Guo K, Zhang SD, Chung T-H, Chng WJ, et al. Metastasis-associated PRL-3 induces EGFR activation and addiction in cancer cells. The Journal of clinical investigation. 2013;123(8):3459.

35. Pathak MK, Dhawan D, Lindner DJ, Borden EC, Farver C, Yi T. Pentamidine Is an Inhibitor of PRL Phosphatases with Anticancer Activity 1 Supported in part by NIH Grants R01CA79891 and R01MG58893 (to TY) and CA90914 (to ECB). 1. Molecular cancer therapeutics. 2002;1(14):1255-64.

36. Min G, Lee S-K, Kim H-N, Han Y-M, Lee R-H, Jeong DG, et al. Rhodanine-based PRL-3 inhibitors blocked the migration and invasion of metastatic cancer cells. Bioorganic & medicinal chemistry letters. 2013;23(13):3769-74.

37. Daouti S, Li W-h, Qian H, Huang K-S, Holmgren J, Levin W, et al. A selective phosphatase of regenerating liver phosphatase inhibitor suppresses tumor cell anchorage-independent growth by a novel mechanism involving p130Cas cleavage. Cancer research. 2008;68(4):1162-9.

38. Salamoun JM, McQueeney KE, Patil K, Geib SJ, Sharlow ER, Lazo JS, et al. Photooxygenation of an amino-thienopyridone yields a more potent PTP4A3 inhibitor. Organic & biomolecular chemistry. 2016;14(27):6398-402.

Submitted: 10/07/17

Review Article

[1]Waksman Institute of Microbiology, Rutgers University, Piscataway, NJ, USA

Keywords

Vibrio natriegens, molecular genetics, bioengineering, fastest doubling time, taxonomy

Email Correspondence

jonah.m.williams@rutgers.edu

Jonah M. Williams[1]

Historical Perspectives on the Bacterium *Vibrio natriegens* and its Potential to Revolutionize Bioengineering

Abstract

Background: *Vibrio natriegens* is an aquatic bacterium that has the fastest doubling time of any currently known organism at approximate 9.8 min. This review delves into the early categorization of *V. natriegens*, its phylogeny and physiology, and the efforts aimed at studying its potential to enhance both micro- and macro-scale biotechnology.

Methods: Twenty-eight research papers from scientific literature databases including PubMED (US National Library of Medicine), National Center for Biotechnology Information (NCBI), and the American Society for Microbiology Journals were used in this study.

Summary: Almost sixty years after the original isolate, microbiologists and bioengineers alike are expressing a renewed interest in *V. natriegens* as a possible replacement for *Escherichia coli*. Recent biotechnological efforts have been successful in developing the necessary genetic systems for such a transition. The productivity of *V. natriegens* suggests that the organism can also be used in large scale bio-refineries producing nutraceuticals and other bio-products.

The Historical Taxonomy of *Vibrio natriegens*

Discovery of the organism and classificiation as a *Pseudomonas*

William Payne first isolated *V. natriegens* from the coastal salt marshes of Sapelo Island, Georgia, USA in 1958. Together with associate R. G. Eagon, Payne studied and categorized the newly discovered bacterium at the University of Georgia in Athens. Originally listed as "marine isolate 11," *V. natriegens* was then classified in 1961 as *Pseudomonas natriegens* due to physiological observations at the time.(1,2) This marine organism is a gram-negative bacterium, the same category as *E. coli* and *Helicobacter pylori*; these bacterium contain periplasmic spaces between their outer and inner membranes.(3) Additionally, *V. natriegens* is quite small, approximately 1 μm in length, and contains a polar flagellum.

Payne and Eagon concluded that the isolated *P. natriegens* required sodium for growth and produced carbon dioxide, acetic acid, lactic acid, and pyruvate from glucose metabolism.(2) Interestingly, when they replaced sodium with other monovalent cations, such as lithium, potassium, and rubidium, in their growth medium, Payne and Eagon observed that the organism did not grow. This experiment validated that sodium was imperative for *P. natriegens* population growth.(4-6) These characteristics refined the organism's taxonomy to either the *Vibrio*, *Pseudomonas*, or *Beneckea* genera.(7, 8) These three genera belong to the Gammaproteobacteria class. Within this class, the *Vibrio* and *Beneckea* genera belong to the *Vibrionaceae* family and the *Pseudomonas* genus belongs to the *Pseudomonadacea* family. Though these families contain rod-shaped, gram-negative bacteria with polar flagella and generally reside in marine/coastal environments, *Pseudomonas* species (spp.) are aerobic while *Vibrio* and *Beneckea* spp. are facultative anaerobes.(7, 9)

At the time of *P. natriegens* classification, the phylogenies of these groups were not well defined. Early observations by Payne et al. (1961) suggested that the isolated organism could belong in any of these genera, but they chose to initially place the organism in *Pseudomonas* after observing no sensitivity to the compound 2,4-diamino-6,7-diisopropyl-pteridine, also known as 0/129.(2) 0/129 is an antimicrobial, vibriostatic agent to which only *Vibrio* bacteria are sensitive and has been used to distinguish *Vibrio* spp. from other gram-negative bacteria.

Improper classification and movement to Beneckea genus

After *V. natriegens*' initial classification into the *Pseudomonas* genus, scientists reclassified the bacterium as existing technologies were developed and improved to better understand its characteristics. In 1971, Baumann et al. proposed in "Taxonomy of marine bacteria: the genus *Beneckea*" that Payne et al.'s genus taxonification introduced in 1961 was inaccurate.(2,7) While categorizing over 145 isolates of marine bacteria which shared many general characteristics with *P. natriegens*, Baumann et al. noted that *P. natriegens* differed in physiology and behaviour from the other existing marine bacteria genera. Based on these findings, Baumann *et al.* reclassified *P. natriegens* into the novel genus *Beneckea*, where all species were gram-negative, had straight rods and polar flagella, required sodium, were unable to fix nitrogen, and had a deoxyribonucleic acid (DNA) Guanine/Cytosine content of 45-48%.(7)

In addition to physiological characteristics, Baumann et al. found that *P. natriegens* was also a facultative anaerobe, a characteristic Payne et al. had also identified. Yet, while earlier analysis showed that carbon dioxide was produced in addition to lactic, pyruvic, and acetic, Baumann et al. found that the organism produced no carbon dioxide through glucose fermentation.(2,7) *Pseudomonas* was a genus for aerobic marine proteobacteria while *Beneckea natriegens* was a gram-negative, facultative anaerobe with a rod-shape and a polar flagellum that fermented glucose without gas production.(13) The differences between Payne et al.'s and Baumann et al.'s experimental observations could have been due to several factors, including different growth conditions, experimental error, strain contamination, captive genetic adaptations, more rigorous experimental analysis, and recent technological advances.(12) These observed experimental differences supported Baumann et al.'s categorization of the bacterium into the *Beneckea* genus.

Reclassification as a *Vibrio* by Austin *et al.* in 1978.

Over the subsequent decade, more scientists grew interested in the phys-

iology of *B. natriegens* and marine proteobacteria. Microbiologists B. Austin, A. Zachary, and R. R. Colwell, revisited *B. natriegens* physiology at the University of Maryland in College Park in 1978 and realized some fundamental flaws in the previous taxonomy classification.(13) They obtained the strain Baumann et al. had been using from the ATCC organismal bank and observed that their original phenotypic analysis was accurate. Austin, Zachary, and Colwell also observed that the organism could not degrade chitin, but could use over 26 other carbon sources, including ethanol, for growth. This result was intriguing, as organisms which can use a wide variety of carbon sources generally have very complex carbon metabolic pathways. *B. natriegens*' ability to synthesize various carbon sources suggested its involvement in biogeochemical carbon cycles and its potential use in artificially biological carbon sequestration and storage initiatives.

Additionally, Austin, Zachary, and Colwell also discovered *B. natriegens* was susceptible to the vibriostatic compound 0/129, a result that contradicted Payne et al.'s original 1961 analysis.(2,13) The different observations on this vibriostatic assay by the two groups of scientists likely results from the nature of the test itself and the available technology in each time period. Researchers performing the same experiment occasionally report varied results, and in the case with *B. natriegens*'s sensitivity to 0/129, the varied mode of chemical preparation, delivery to the organism, and growth conditions all impact the detection ability. Albeit, the specific vibriostatic assay information was in neither Baumann et al.'s nor Payne et al.'s manuscripts. Nonetheless, *B. natriegens*' susceptibility to 0/129 and its phenotypic similarities with other *Vibrio* spp. prompted Austin et al. to suggest that *B. natriegens* be recategorized under the *Vibrio* genus. New evidence published by Payne in 1971 further corroborated Austin et al.'s reclassification proposal. In 1978, *B. natriegens* was renamed as *Vibrio natrigens.*(4,13)

Breaking a Scientific Record

In the early 1960s, Eagon and Payne continued their work with *Vibrio natriegens.* Eagon himself published "*Pseudomonas natriegens*, A Marine Bacteria with a Generation Time of Less than 10 Minutes," in 1962.(14) Eagon reported a record generation time of 9.8 min. for *V. natriegens* under optimal growth conditions in a brain-heart infusion broth at 37^0C with 1.5% sea salt.(14) The observed 9.8 min. generation time of *V. natriegens* was significant as the model organism at the time, *E. coli*, had a laboratory doubling time of approximately 25-30 min.(15, 16) Given how bacteria must increase their volume, replicate and proofread their DNA, synthesize proteins, and undergo cytokinesis, it is truly remarkable that *V. natriegens* can double its population in under 10 min.(17) The organism's generation time expands our limited understanding of the kinetic and thermodynamic mechanisms of cell division and organelle biogenesis.(17,18)

Genomics in Explaining Doubling Time

With recent advances in methods to quickly analyze genomes and transcriptomes, researchers are now working to understand the biochemical mechanisms for *V. natriegens*' doubling time. In 2002, Aiyar et al. attributed the bacterium's extremely fast generation time to very rapid protein synthesis rates.(19) They hypothesized that *V. natriegens* has both a higher count of ribosomes and stronger ribosomal activity mediated by powerful ribosomal RNA (rRNA) promoters and operons. Their results showed that *E. coli* produced approximately 70,000 ribosomes per cell within its 25 min. doubling time, whereas *V. natriegens* produced 115,000 ribosomes per cell within 10 min.(19) relative to a similar average cell size of about 1 micron. Through southern blotting, Aiyar et al. also identified about 13 rRNA operons in *V. natriegens* and found that the promoters of these operons were highly regulated and heavily reliant upon upstream regulators. To determine the strength of these promotors, Aiyar et al. transformed the rRNA production systems into *E. coli* and compared the protein production levels in vitro.(19)

Although knowledge of *V. natriegens* ribosomal RNA promoters and operons behaviour is a significant step towards explaining its replication rate, no genome of *V. natriegens* was available at that time. Maida et al. addressed this problem and submitted a draft genome sequence of *V. natriegens* in 2013.(20) The sequence allowed cross-referencing of *V. natriegens*' observed physical data to the reported DNA sequence and provided a comprehensive link between physiology and genetics. Maida et al. performed a physiological assessment of the bacteria under similarly ideal conditions to Eagon's 1962 experiment and noted that under strong aeration, increasing nutrient supply via mass transfer, the *V. natriegens* culture reduced its doubling time to 7 min.(14,20) With further analysis with homology search bioinformatics, the researchers also discovered 14 rRNA-encoding genes and were able to predict a total of 12 putative rRNA operons. This result aligned closely with the number of rRNA operons discovered by Aiyar et al. a decade prior.(19,20) Additionally, Maida et al.'s genomic sequencing identified closely associated rRNA promotors, information crucial for developing *V. natriegens* expression systems and genetic transformations for increasing biomass yields or productivities.

Like researchers before them, Lee *et al.* discovered 11 rRNA operons within the genome and identified 129 transfer RNA (tRNA) genes. In addition, Lee et al. provided the first complete and fully annotated genome of *V. natriegens* with improved genetic spatial organization.(22) These new findings may have implications on the doubling time of *V. natriegens* as aforementioned works have suggested that an increased number of rRNA operons and available tRNA generally results in faster doubling times.(22) More ribosomes and tRNAs increase a cell's translational capacity, enabling the bacterium to produce more functional proteins in a shorter time period. *V. natriegens* has 4 more rRNA operons and 30 more tRNA genes than *E. coli*, which could explain *V. natriegens*' impressive doubling time.(22) However, as noted by Aiyar et al., more rRNA genes cannot causally imply a faster generation time as the lack of transcriptomic information for both *E. coli* and *V. natriegens* prevents exact genomic comparisons.(19) Furthermore, limited data on translational regulation, ribosome assembly, and rRNA operon regulation has made determining mechanisms underlying *V. natriegens*' rapid doubling time difficult.

Assessing the Biotechnological Potential of V. natriegens

Since Eagon's initial categorization of *V. natriegens* in the 1950s and 60s, much of the research focus in following decades had been on the physiology explaining *V. natriegens*' growth capacity. Yet renewed interest in the organism stems mainly from a desire to apply its fast doubling time to biotechnological experiments. The use of *V. natriegens* as a surrogate organism could impact a variety of professionals, from research scientists looking for fast host-vector DNA/protein expression systems to engineers operating scaled bioreactors to mass-produce valuable bio-commodities. However, for *V. natriegens* to replace the current biological standard, *E. coli*, its full biotechnological profile needs to be developed and extensively studied, a project which many research groups have begun to undertake. Two champions in the area of *V. natriegens* biotechnology are George Church et al. at Harvard University in Cambridge, MA and members from Synthetic Genomics Inc.(SGI) in La Jolla, CA. Both research teams published papers on integrating *V. natriegens* into key biotechnological systems nearly a month apart during the summer of 2016. They presented almost identical research approaches and conclusions and highlighted the desired functionalities in a model benchtop organism.(22,23)

Researchers have also recently developed efficient transformation and recovery protocols for delivering plasmids into *V. natriegens* to yield

between 105-107 CFU/μg recombinant DNA per reaction.(22,23) *V. natriegens*' transformation efficiency is similar to that of *E. coli*; this illustrates the closeness between these two organisms as genetic surrogates. Additionally, bacterial conjugation from *E. coli* to *V. natriegens* is possible, whereby the plasmids were maintained as episomal bodies inside the latter, allowing scientists to integrate DNA from more advanced organisms and develop co-cultures to produce valuable bio-products. (23,24) It was also shown that while transformed *V. natriegens* colonies appeared after 5-6 hr, wild varieties grown on agar plates had observable colonies after only 4 hr, nearly 2.5 times faster than a similar *E. coli* system grown in rich media.(22) Weinstock et al. showed successful maintenance of antibiotic resistance cassettes (such as ampicillin, kanamycin, and chloramphenicol) within the organism and discovered that the lacI, araBAD, and λ phage pR inducible promotors function in the organism. (23) These resistance and expression systems are key components in molecular biology laboratories, and their presence in *V. natriegens* furthers its ability to compete with *E. coli* in transformation, genetic cloning, and mutant generation capabilities.

Determining optimal growth conditions, such as nutrients, temperature, and concentration, strongly impact the costs associated with efficient and beneficial energy sources. Organisms which can use a wide variety of carbon sources are inherently advantageous. In their recent publications, the Harvard and SGI teams observed that in minimal media, *V. natriegens* thrived on many carbon sources, including sucrose. Lee et al. from Harvard University reported that the organism reached the stationary growth phase after only 6-7 hr in liquid media supplemented with sucrose.(22) In comparison, *E. coli* is unable to use sucrose, a relatively inexpensive sugar, as a carbon source. Weinstock et al. from SGI were able to develop robust T7 RNA polymerase expression systems in *V. natriegens* identical to those currently found in *E. coli* and could potentially both express and recover GFP under an inducible IPTG promotor.(23) Weinstock et al. used SDS-PAGE analysis and fluorescence measurements from GFP to quantify these results.(23) Cre-Lox recombination was successfully performed for *V. natriegens* and Lee et al. were also able to develop a transposon mutagenesis system in addition to a basic CRISPRi gene regulation system.(22-24) Cre-Lox, CRISPRi, and protein expression systems are all fundamental genomic editing and analysis techniques for any organism involved in transgenic investigations, personalized medicine, mutagenesis, protein engineering, etc. The rapid development of gene and protein expression techniques and the ability of *V. natriegens* to undergo genomic engineering similar to *E. coli* establishes *V. natriegens* as a promising laboratory surrogate.

Uprooting the Status Quo

Although the baseline that technology scientists could use to substitute *V. natriegens* for other surrogate organisms in the lab is now available, it is uncertain if researchers will switch to *V. natriegens*. Even with its impressive doubling time, *V. natriegens* increases work efficiency marginally in the laboratory setting. Transformed *V. natriegens* colonies appear on solid media after approximately 5.5 hr, whereas *E. coli* colonies under the same conditions would take about 12 hr to appear.(23) If a researcher is extremely pressed for time and needs to rapidly collect data, perhaps using *V. natriegens* for one experiment would save time. Even so, the amount of time thsat *V. natriegens* can save is arguably marginal considering the resources expended for developing novel protocols and introducing a new organism to the laboratory. The inconvenience of replacing *E. coli*, which has been the gold-standard for biologists for decades, with *V. natriegens* is therefore extremely challenging.(15)

E. Coli is a universal model organism and is commercially available in a wide variety of strains, auxotrophs, and mutants. *E. coli* has been intensively studied since 1885 when Theodor Escherich discovered the bacterium.(15) Since the 1940s, non-pathogenic forms of *E. coli* have been used in experiments to resolve key aspects of prokaryotic physiology and behaviour. As time progressed, our understanding of fundamental molecular biological processes grew alongside developments in the *E. coli* system, further reinforcing its use in molecular and cellular biology.(25-27) Categorized, resistant, and harmless, *E. coli* is a manageable bacterium that replicates quickly. With each new protein theorized, gene discovered, or pathway of interest determined, *E. coli* was, and still is, the first organism used to develop an expression system, study metabolism, or generate mutations to illustrate molecular processes. Our comprehension of *E. coli* "omics" (genome, proteome, transcriptome, lipidome, interactome, etc.) is as extensive as the global laboratory use of the organism and has allowed researchers to understand the true effects of any mutation or genetic engineering.(25, 26) Functionally, both *E. coli* and *V. natriegens* are very similar: they are generally the same size, have flagella, and are gram-negative. Yet despite their close relationship, it appears that the widespread availability, functionality, cost effectiveness, and traditions surrounding *E. coli* makes it unlikely that *V. natriegens* will replace *E. coli* in the near-future.

Shortly after the SGI team described *V. natriegens* gene and protein expression systems, the company commercialized VmaxTM Express Competent cells, an electroporation-ready strain of *V. natriegens* equipped with an IPTG-inducible T7 promotor system and a doubling time of 14 min.(23) Vmax now provides scientists with a new choice for molecular cloning and expression studies and can potentially transform bioengineering. As a protein or metabolite expression and production system, *V. natriegens* could greatly improve large-scale yields and operation times. Biopharmaceutical companies use surrogate organisms to produce a number of small-molecule and protein-based drugs in large bioreactors, and *V. natriegens*' faster replication time can substantially increase biomass production, and therefore drug or bioproduct.(26, 27, 28) Although there may not be many cost benefits to using *V. natriegens* at the benchtop level, cost reductions may be realised in large-scale bio-production operations.(28) *V. natriegens* can benefit large-scale bio-refineries looking to cut costs and enhance productivity without compromising quality.

Decreasing the costs and time for bio-product assembly can also benefit the development of personalized medicines.(26, 27) For example, *V. natriegens* is promising in pharmacogenomics and for developing drugs that address genetic based disorders and deficiencies. *V. natriegens*' rapid growth and stability as a genetic vector enhances genomic screening of medical conditions. Through the use of protein expression systems or RNAi/microRNA/CRISPRi, researchers can advance treatments to silence harmful genes, introduce necessary exogenous proteins, and provide faster treatments for patients with novel or uncommon conditions. Perhaps one the most interesting characteristics of *V. natriegens* applicable to the biomedical field is its remarkable ability to secrete proteins directly into the growth medium. As noted by Weinstock and the Synthetic Genomics team, this capability will allow for easier bio-separation processes, thus saving costs downstream.(23)

Conclusion

As for *V. natriegens* itself, much is left to learn about the physiology and mechanisms behind its doubling time. Further understanding its impressive growth phenotype and physiology can uncover genetic and biochemical systems that could increase biofuel crop growth, fight its pathogenic relative *V. cholerae*, or advance understanding of the salt marsh microbiome. For now, a general lack of attention and behavioural description of *V. natriegens* limits the bacterium's adoption in bio-production. Regardless, the biotechnological future looks propitious for *V. natriegens* as scientists have only scratched the surface of this robust microbe.

References

1. Payne, WJ. Studies on Bacterial Utilization of Uronic Acids III. J Bateriol. 1958;76:301-307. PMCID: PMC290205
2. Payne WJ, Eagon RG, Williams AK. Some observations on the physiology of Pseudomonas natriegens noc. spec. Antoine Van Leeuwenhoek. 1961;27:121-128.
3. Silhavy TJ, Kahne D, Walker S. The bacterial cell envelope. Cold Spring Harb Perspect Biol. 2010;2(a000414). doi:10.1101/cshperspect.a000414.
4. Webb CD, Payne WJ. Influence of Na+ on synthesis of macromolecules by a marine bacteria. Appl Microbiol. 1971;21:1080-1088.
5. MacLeod R. The question of the existence of specific marine bacteria. Bacteriol Rev. 1965;29:9-23.
6. Oh S, Kogure, Ohwada K, Simidu U. Correlations between possession of a respiration-dependent Na+ pump and Na+ requirement for growth of marine bacteria. Appl Environ Microbiol. 1965;57:1844-1846.
7. Baumann P, Baumann L, Mandel M. Taxonomy of marine bacteria: the genus Beneckea. J Bacteriol. 1971;107:268-294.
8. Dikow RB. Systematic relationships within the Vibrionaceae (Bacteria: Gammaproteobacteria): steps toward a phylogenetic taxonomy. Cladistics. 2011;27:9-28. doi:10.1111/j.1096-0031.2010.00312.x.
9. Farmer JJ. The family Vibrionaceae. In: Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E, editors. The Prokaryotes. 3rd ed. Singapore: Springer; 1965. p. 495-508.
10. Shewan JM, Hodgkiss W, Liston J. A method for the rapid differentiation of certain non-pathogenic asporogenous Bacilli. Nature. 1954;173:208-209.
11. Matsushita S, Kudoh Y, Ohashi M. Transferable resistance to the vibriostatic agent 2,4-diamino-6,7-diisopropyl-pteridine (0/129) in Vibrio cholerae. Microbiol Immun. 1984;28:1159-1162. doi:10.1111/j.1348-0421.1984.tb00773.x
12. Beaume M, Monina N, Schrenzel J, François P. Bacterial genome evolution within a clonal population: From in vitro investigations to in vivo observations. Future Microbiol. 2013;8:661-74.
13. Austin B, Zachary A, Colwell RR. Recognition of Beneckea natriegens (Payne et al.) Baumann et al. as a member of the Genus Vibrio, as previously proposed by Webb and Payne. Int J Syst Bacteriol. 1978;28:315-317.
14. Eagon RG. Pseudomonas natriegens, a marine bacteria with a generation time of less than 10 minutes. J Bacteriol. 1962;83:736-737.
15. Blount ZD. The unexhausted potential of E. coli. Elife. 2015 Mar 25;4. doi: 10.7554/eLife.05826.
16. Studier W, Daegelen P, Lenski RE, Maslov S, Kim J. Understanding the Differences between Genome Sequences of Escherichia coli B Strains REL606 and BL21(DE3) and Comparison of the E. coli B and K-12 Genomes. J. Mol. Biol. 2009;394:653-680.
17. Wang JD, Levin PA. Metabolism, cell growth and the bacterial cell cycle. Nat Rev Micobiol. 2009;7:822-827. doi:10.1038/nrmicro2202.
18. Iyer-Biswas S, Wright CS, Henry JT, Lo K, Burov S, Lin Y, et al. Scaling laws governing stochastic growth and division of single bacterial cells. 2014;111:15912-15917. doi:10.1073/pnas.1403232111.
19. Aiyar SE, Gaal T, Gourse RL. rRNA promotor activity in the fast-growing bacteria Vibrio natriegens. J Bacteriol. 2002;184:1349-1358. doi:10.1128/JB.184.5.1349-1358.2002.
20. Maida I, Bosi E, Perrin E, Papaleo MC, Orlandini V, Fondi M, et al. Draft genome sequence of the fast-growing bacteria Vibrio natriegens strain DSMZ 759. Genome Announc. 2013;1:e00648-13. doi:10.1128/genomeA.00648-13.
21. Wang Z, Lin B, Hervey WJ, Vora GJ. Draft Genome Sequence of the Fast-Growing Marine Bacteria Vibrio natriegens Strain ATCC 14048. Genome Announc. 2013 Aug 8;1.(4) pii: e00589-13. doi: 10.1128/genomeA.00589-13.
22. Lee HH, Ostrov N, Wong BG, Gold MA, Khalil AS, Church GM. Vibrio natriegens, a new genomic powerhouse. 12 June 2016, posting date bioRxiv 058487 doi:https://doi.org/10.1101/058487
23. Weinstock MT, Hesek ED, Wilson CM, Gibson DG. Vibrio natriegens as a fast-growing host for molecular biology. Nature Methods. 2016;13:849-851. doi:10.1038/nmeth.3970.
24. Lee HH, Ostrov N, Gold MA, Church GM. Recombineering in Vibrio natriegens. 26 April 2017, posting date. bioRxiv 130088 doi:https://doi.org/10.1101/130088
25. Zhang H, Fang L, Osburne MS, Pfeifer BA. The continued development of E. coli as a heterologous host for complex natural product biosynthesis. Methods Mol Biol. 2016;1401:121-134. doi:10.1007/978-1-4939-3375-4_8.
26. Baeshen MN, Al-Hejin AM, Bora RS, Ahmed MM, Ramadan HA, Saini KS, et al. Production of biopharmaceuticals in E. coli: current scenario and future perspectives. J Microbiol Biotechnol. 2015;25:953-962. doi:10.4014/jmb.1412.12079.
27. Mahalik S, Sharma AK, Mukherjee KJ. Genome engineering for improved recombinant protein expression in Escherichia coli. Microbial Cell Factories 2014;13:177. doi:10.1186/s12934-014-0177-1.
28. Tripathi NK. Production and purification of recombinant proteins from Escherichia coli. ChemBioEng Rev. 2016;3:116-133. doi:10.1002/cben.201600002